Spring 1943

The Standardization of The Kelly Spatial Insight Test

Darlieene Thompson
Fort Hays Kansas State College

THE STANDARDIZATION OF THE KELLY SPATIAL INSIGHT TEST

being

A Thesis presented to the Graduate Faculty of the

Fort Hays Kansas State College

in partial fulfillment of the requirements for

the degree of

Master of Science

by

Darlienne Thompson, B. S. in Education

Fort Hays Kansas State College

Approved: _____
Major Professor

3-11-'43
date

_____
Chrmn. Graduate Council

TABLE OF CONTENTS

LIST OF TABLES

LIST OF CHARTS

LIST OF DIAGRAMS

INTRODUCTION

(Previous work in the field of performance tests): In the measurement of intelligence, one has frequently called to his attention certain factors which affect the score on the intelligence test but which do not seem properly to belong to intelligence. Binet found it necessary to rule out of intelligence certain information which was not equally available to all children. Pintner and Paterson (9) attempted to rule out verbal facility when theyconstructed their performance scale twenty-four years ago. Following the publication of the Army Alpha intelligence test, some psychologists made a more determined attempt to rule out such information as was contained in a large section of that test although other psychologists, more intent upon impressive numbers of administrations than upon validity, continued to place reliance upon the Army Alpha. In recent years an attempt has been made to rule out the factor of serial reaction time and we find some tests constructed with an administration time sufficient to allow ninety-five per cent of the subjects to proceed at their natural rate. Two discoveries made this step seem necessary: The discovery that the Army Alpha test correlated highly with serial reaction time, (10) and the discovery that one may increase the reliability coefficient of test indefinitely simply by shortening the time limit. (10)

It seems that if the intelligence test is to become practically valuable it should be constructed to give the examiner some basis for predicting the intelligence of a person's behavior under

various conditions. Thus, in a school with conventional curriculum, the ordinary type of academic intelligence test might be valuable, but in a clinic, where a wide variety of different educational devices are available, many factors which affect the intelligence score should be singled out and measured separately. Obviously the following factors require special consideration: Sensory disabilities; physical disabilities, including fatigue, incoordination, and irregular energy output; discrepancies in educational opportunity and social background; unequal motivation and personal attitudes toward the past; distractibility; and speech disorders. Additional factors which clinical experience would lead us to believe that we are now ready to consider separately from the main study of intelligence are the aphasic disorders such as apraxia, auditory aphasia, visual aphasia, alexia, agraphia, transient aphasic conditions, etc. These disorders are only a step removed from the physical sensory and motor disorders which nearly everyone would agree should be considered separately from the rest of intelligence (10).

The performance tests are important for the following purposes:

1. To test "verbalists" (the subnormal) who are likely to show too great an advantage in conversational tests.

2. To test children of "first-rate intelligence" who cannot express themselves clearly at any length on paper and who fail to do themselves justice in verbal answers to verbal questions.

3. To test special aptitudes for selection of pupils for junior technical schools. (4)

By the means of well-selected performance tests we can measure a factor of fairly wide application, which has some right to be considered a part of general intelligence. Many studies have shown the relatively independent character of tests of manual dexterity, mechanical ability, and certain specific abilities such as formboard ability, in reference to "general intelligence" as measured by verbal tests or by school success (4). The increasing emphasis on the relative independence of the group factors underlying linguistic ability, arithmetical ability, mechanical ability, and social ability, which has been made by various authors with various modes of approach, suggests that in our attempt to predict general mental adaptability, performance tests have a real function to perform. Cornell (4) and Coxe (4) list the functions of performance tests as follows:

1. To offer standardized conditions for observing the behavior of an individual in more varied situations than are provided by verbal tests.

2. To prevent the drawing of too sweeping conclusions from observation based on a narrow range of material.

3. To test the hypothesis that there is a group factor underlying general concrete ability, which is of importance in the concept of general intelligence.

4. To utilize for purposes of guidance specific kinds of tests, such as tests of mechanical ability, of motor dexterity and control, and of other special abilities (4).

(Development of the Pintner-Paterson Scale): The work of developing a group of tests which did not involve any kind of language response was begun in 1914 with the standardization of a few performance tests. In 1917, Pintner (9), then

assistant professor of psychology at Ohio State University, and Paterson (9), then an instructor in psychology at the University of Kansas, constructed the Pintner-Paterson Scale. Several tests used by Knox and Healey were incorporated into the scale and have proved to be very valuable. The tests devised by Knox that are included in the Pintner-Paterson Scale are number five, the Casuist Form Board and number ten, the Feature-Profile Test. Knox colaborated with Kempf to make this test, and the Knox Cube Test which was modified by Pintner. Healey's contribution to the scale consisted of the Mare and Foal Picture Board, test number eleven, which was modified before it was included in the scale, the Healey Construction Puzzle A, test number eight, and the Picture Completion Test, test number twelve. Other psychologists who contributed to the scale are Twitmeyer, Goddard, Gwyn, Kempf, Glueck, Woodworth, and Wells. The scale is especially for the groups of children who are handicapped by language disabilities such as the foreign child, the speech defective, and the deaf child. A scale of performance tests is the only adequate means for the measurement of mentality for these groups of children (9). The language factor must be omitted and the estimate of mentality must be based upon what any of these kinds of children can do as compared with the normal hearing and speaking child. The Pintner-Paterson rules out the verbal ability but is greatly influenced by motor disabilities and poor muscular coordination. The problems are simple to begin with and the more difficult tests are merely more complex tasks, that is, they are complicated simple tasks. It does not attempt to test a higher task level and it is not satis-

factory for testing higher levels of intelligence.

Each board was standardized separately, part of the time the standardization of the original author is accepted and incorporated into the scale.

(Development of the Cornell-Coxe Performance Ability Scale): In the Cornell-Coxe Performance Ability Scale (4) there are seven separate tests that may be given in any order. This Scale was copyrighted in 1934 by the World Book Company. The authors, Cornell, then a psychologist in the New York State Education Department, and Coxe, then director of research in the New York State Education Department, selected these particular seven tests partly upon the basis of previous experience and partly upon the basis of experimentation. The seven tests are:

1. Manikin-Profile: Same test that is used in the Pintner-Paterson Scale and also in the Army Performance Scale. This test has a correlation of $r = .73 \pm .02$ with the total performance score.

2. Block-Design: This was adapted from Kohs' test but the method of scoring was changed. It has a correlation of $r = .86 \pm .01$ with the total performance score.

3. Picture-Arrangement: This was devised by the authors from suggestions from the works of Decroly and Vermeylen. A correlation of $r = .84 \pm .02$ with the total performance score was found.

4. Digit-Symbol: This is the same form that was used in the Army Performance Scale. Its' correlation with total performance score is $r = .83 \pm .01$.

5. Memory-for-Designs: This is a modification of the test used in the Army Performance Scale. The correlation with the total performance score is $r = .81 \pm .01$.

6. Cube-Construction: This is also from the Army Performance Scale but the method of scoring is different. Correlation with the total performance score is $r = .56 \pm .05$.

7. Picture-Completion: Healy's Pictorial Completion Test II. Test number three and test number seven are intended to be used interchangeably, so that the series as given consists of six tests. Correlation with the total performance score is r = .77 ± .02.

The correlation of the total performance score with chronological age was r = .78.

In one of the schools used in the standardization, the test was repeated on 125 children after an interval of eleven months. The correlation between the first and the second testing was r = .92. In view of the time interval this seems very satisfactory. The probable error of estimating the second score from the first, on the basis of this correlation, is 12.2 points of score or approximately eight months of mental age. This is approximately six per cent of the total range of the scores obtained in the first testing and about eight per cent of the median. This may seem large but it is no larger than the PE of the Stanford-Binet scale as found by Otis. The reliability of each of the separate tests is shown by the following correlations:

Manikin-Profile	r = .66	Memory-for-Designs	r = .81
Block-Design	r = .78	Cube-Construction	r = .76
Digit-Symbol	r = .89	Picture-Completion	r = .70

The norms given for the test are given in age norms. The Cornell-Coxe (4) has a higher task level but it does not recognize the difference in task levels. That is, the test may be given in any order and not in an order from simple tasks to a higher task level.

(Development of the Terman-Merrill Revision of the Stanford-Binet): In 1916, Terman wrote "The Measurement of Intelligence" (11) which was an explanation of and a complete guide for the use of the

Stanford Revision and extension of the Binet-Simon Intelligence Scale. The Binet scale is made up on an extended series of tests in the nature of "stunts" or problems, success in which demands the exercise of intelligence. As left by Binet, the scale consists of 54 tests, so graded in difficulty that the easiest lie well within the range of normal 3-year-old children, while the hardest tax the intelligence of the average adult. The problems are designed primarily to test native intelligence, not school knowledge or home training. In 1937, Terman, then professor of psychology at Stanford University, and Merril, then associate professor of psychology at Stanford University, contributed to the field of performance tests when they published their complete revision and extension of the original Stanford-Binet scales (12) basing the revision and standardization upon larger and more representative groups. The major faults of the original Stanford-Binet scale have long been recognized. Although affording a satisfactorily valid and reliable measure over a fairly wide intermediate range, it was especially defective at both extremes. Abilities below the mental level of four years or above that of the average adult were very inadequately sampled. One of the severest limitations to the usefulness of the scale was the fact that no alternative form was available for use in retesting or as a safeguard against coaching. In the revision two scales were provided which differ in content but are mutually equivalent with respect to difficulty, range, reliability, and validity. The scales are designated as Form L and Form M. In content, Form L bears greater resemblance to the original Stanford-Binet, but neither form can be recom-

mended above the other. As was said before the scale devised by Binet contained 54 tests, and the first Stanford revision increased the number to 90. Each form of the new revision contains 129 tests. Below the five-year level tests are located at half-year intervals. The scale makes use of diminutive objects, brightly colored cubes, wooden beads and other attractive material. In general, the content of the new scales resembles the old scale including such tests as comprehension, absurdities, word-naming, drawing designs, memory-for-digits, giving differences and similarities, and defining abstract terms. The fact that the same subjects were used in the standardization of Form L and Form M has made it possible to guarantee almost perfect equivalence of the scores yielded by the two tests. Average schools were chosen and the pre-school group were recruited from the siblings of school cases. All subjects are American-born and belong to the white race. There has been no elimination of any particular nationality groups. The majority of the children aged four and over were examined in an unused room at the neighborhood school building where they were brought either by the mother or by the examiner who called for and returned children whom the mother was unable to bring. The majority of the children under four were tested in their own homes. The reliability values range from .98 for subjects below 70 I. Q. to approximately .90 for subjects above 130 I. Q. For subjects near 100 I. Q. the reliability is .925. The reliability values are as follows:

I. Q. 130 and over	r = .898
I. Q. 110-129	r = .912
I. Q. 90-109	r = .924
I. Q. 70-89	r = .945
I. Q. Below 70	r = .982

The reliabilities found by correlating Form L against Form M ranged from .85 to .95 with a madian of .91.

This is a little of the previous work in the field of performance tests. The work of the author has been confined mainly to standardization of a group of tests already constructed by Kelly, Director of the Psychological Clinic at Fort Hays Kansas State College.

STATEMENT OF PROBLEM

The purpose of this study is to standardize the Kelly Spatial Insight Test, a test of performance ability. The task was to establish norms by which individual cases can be compared with each other.

The use of form boards in the measurement of intellectual performance should be accompanied with considerable caution because (1) only a relatively small area of performance is involved, (2) motor handicaps influence the score so markedly, and (3) most boards may be filled without use of any intellectual observation more crucial than the direct comparison of angles, lengths of lines, areas, and other incidental details. Proper caution of the first two points is simply a matter of good clinical procedure; the proper use of caution on the third point involves the construction of better form boards.

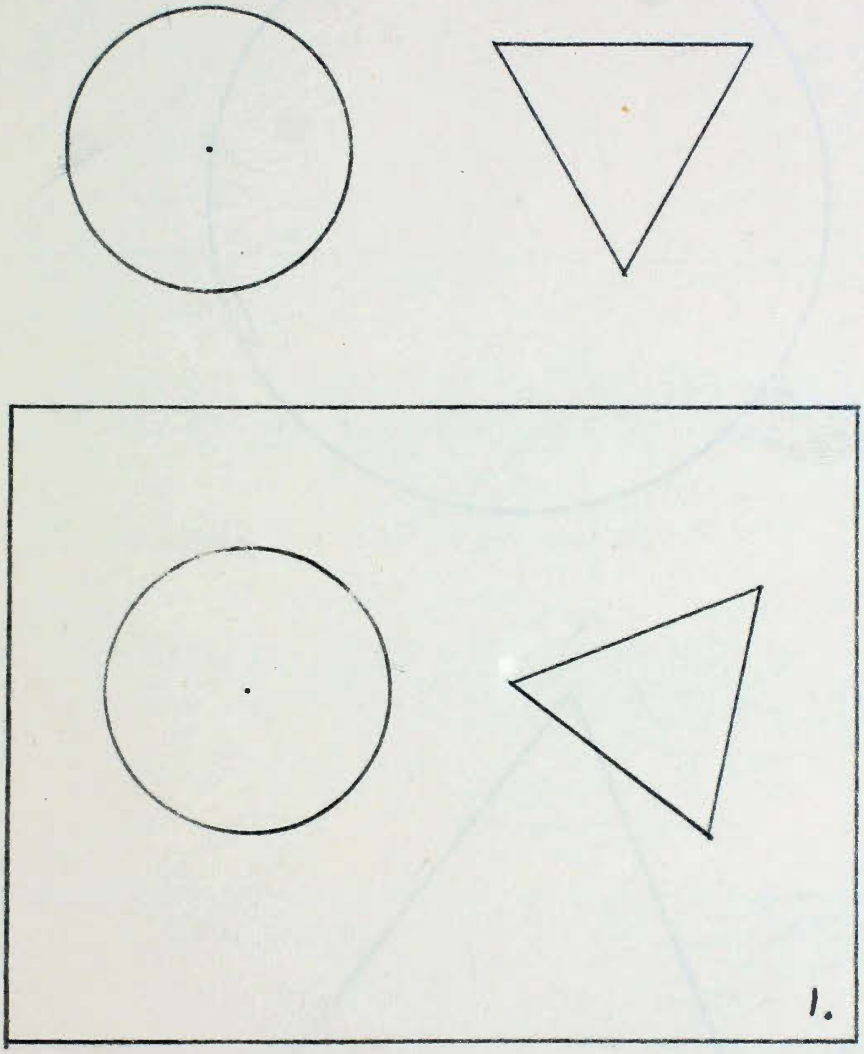
With these points in mind Kelly attempted to construct a scale of form boards for use with subjects between the ages of eighteen months and adulthood. In the lower part of the scale the boards are constructed according to conventional patterns and require for their

board cut to carry the extra inserts necessary for the test. This board is made on the same order as those of tests 1-7. Two nails are included with each set of tests 10, 11, and 12. The size of the nails is not of consideration except they must be easily inserted through the correct holes in the inserts. Each board is numbered in black ink in the lower right-hand corner. The extra board for tests 10, 11, and 12 are numbered the same as the test with a capital "A" following the number. The boards are designed to be given in that order. The correct pieces fit easily into the correct recess and in the upper range the function when correctly put together moves easily for the subject. The entire test fits into a wooden box approximately 12½" long, 9½" wide, and 8½" high.

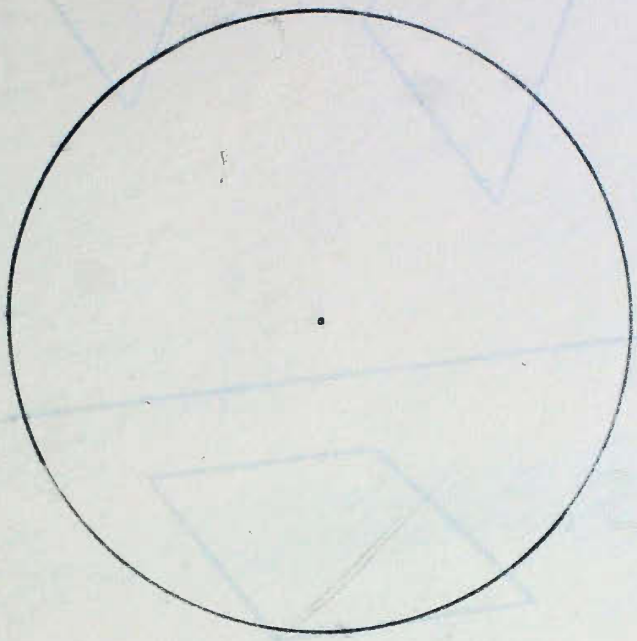
(Description of individual boards): Each board is named according to the type of insight required to solve it.

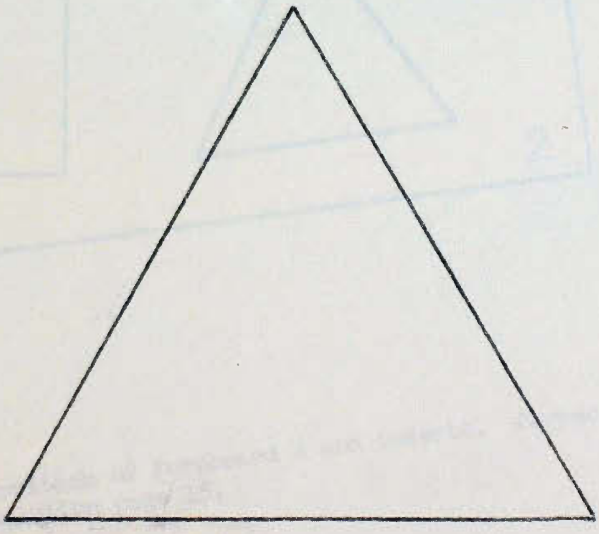
1. Gross Shape Discrimination: A circle and a triangle, two dissimilar forms. The circle and triangle are painted white. The circle is 3 5/8" in diameter and the triangle is an equilateral triangle each side being 3 3/8" long. The recesses are slightly larger so the pieces slip into them very easily. One angle of the triangle points towards the circle and no side is parallel with the edge of the form board. The test is to make a discrimination between gross shapes.

2. Angle: A triangle, a diamond, and a square set in the board in such a relation to each other and to the margins of the board as will accent their similarities. The three inserts are painted white. Each side of the square is 2 5/8", each side of the diamond is 2 5/8" and the triangle is 3 3/8" along each side. As the board faces the

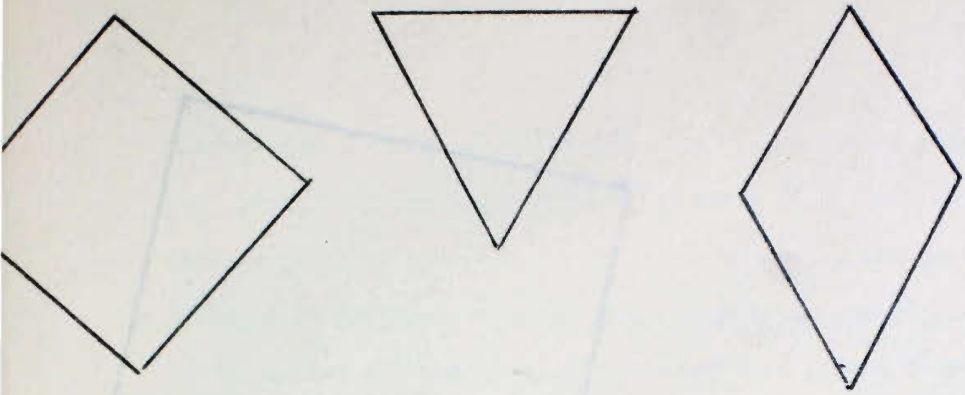


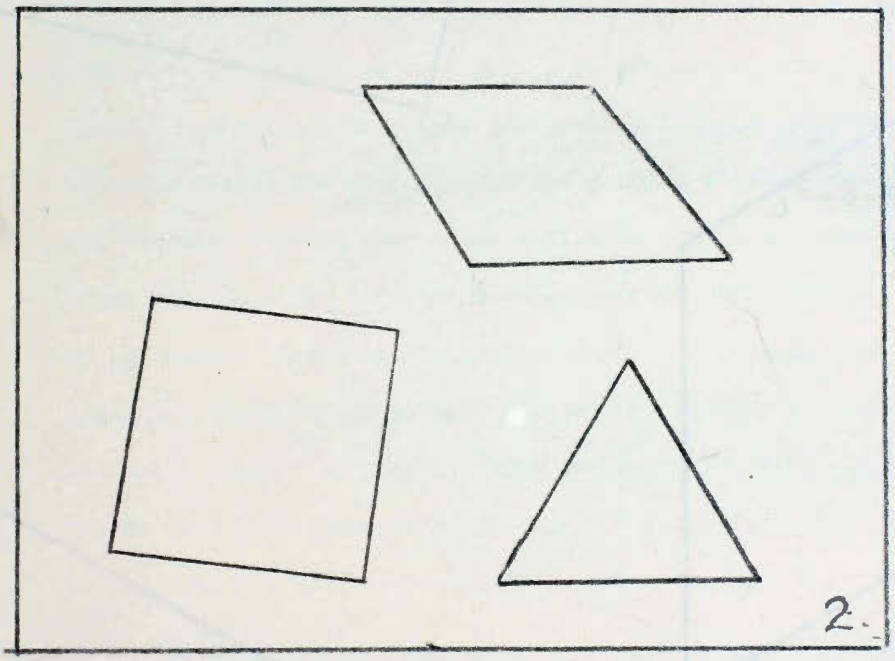
Position of Formboard 1 and inserts for first trial. Correct solution page 13.



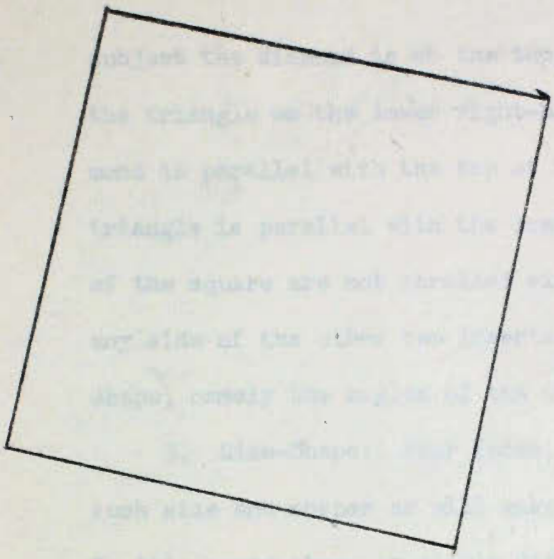


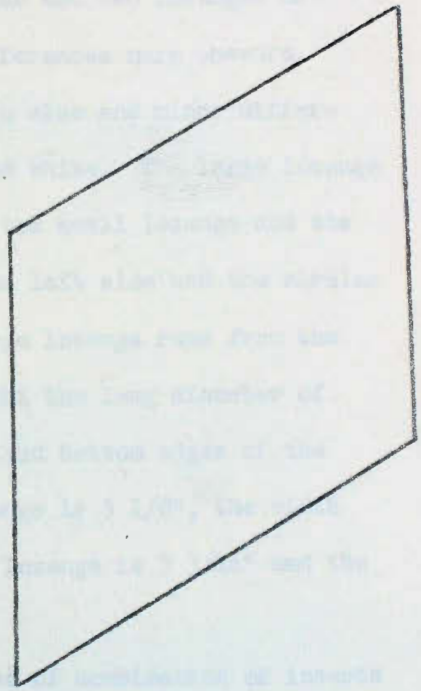
Correct Solution

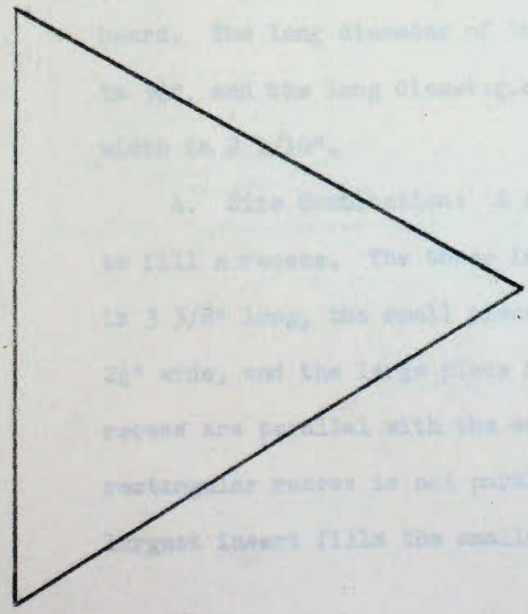




Position of Formboard 2 and inserts. Correct solution page 15.





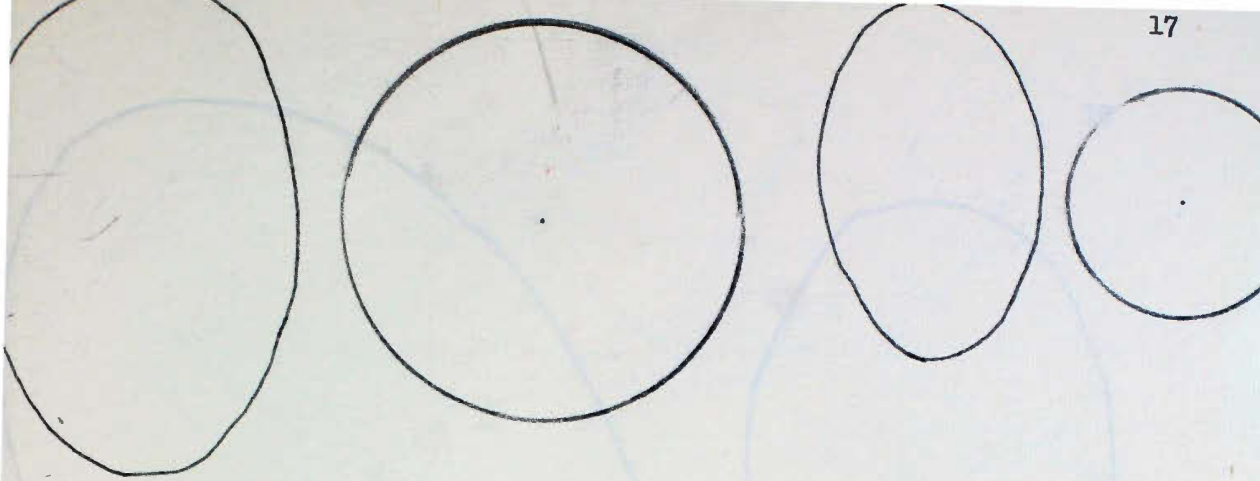


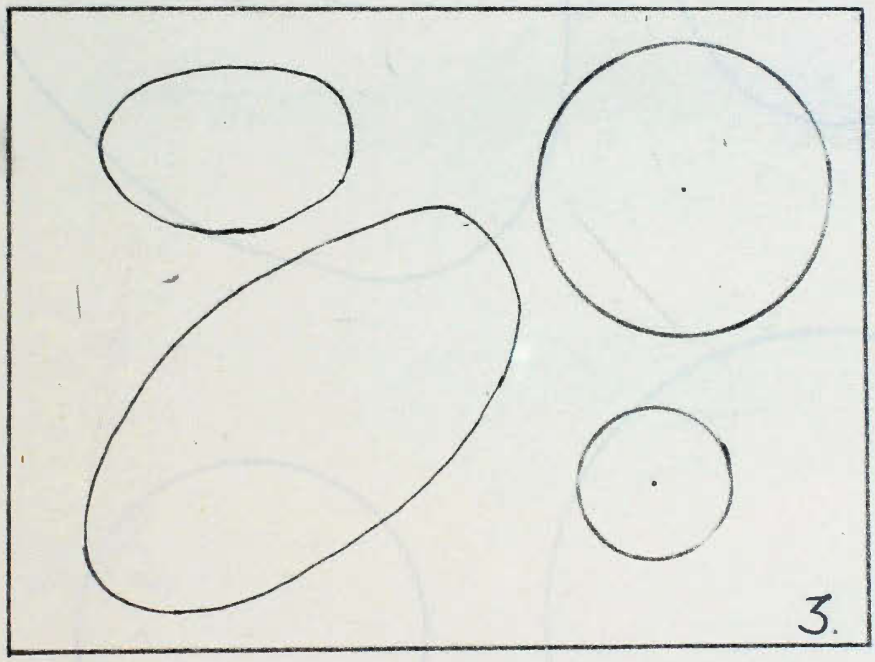
Correct Solution

subject the diamond is at the top, the square on the lower left and the triangle on the lower right-hand side. The top side of the diamond is parallel with the top of the form board, the base of the triangle is parallel with the lower edge of the board but the sides of the square are not parallel with any edge of the form board or any side of the other two inserts. The test is one concerned with shape, namely the angles of the three inserts.

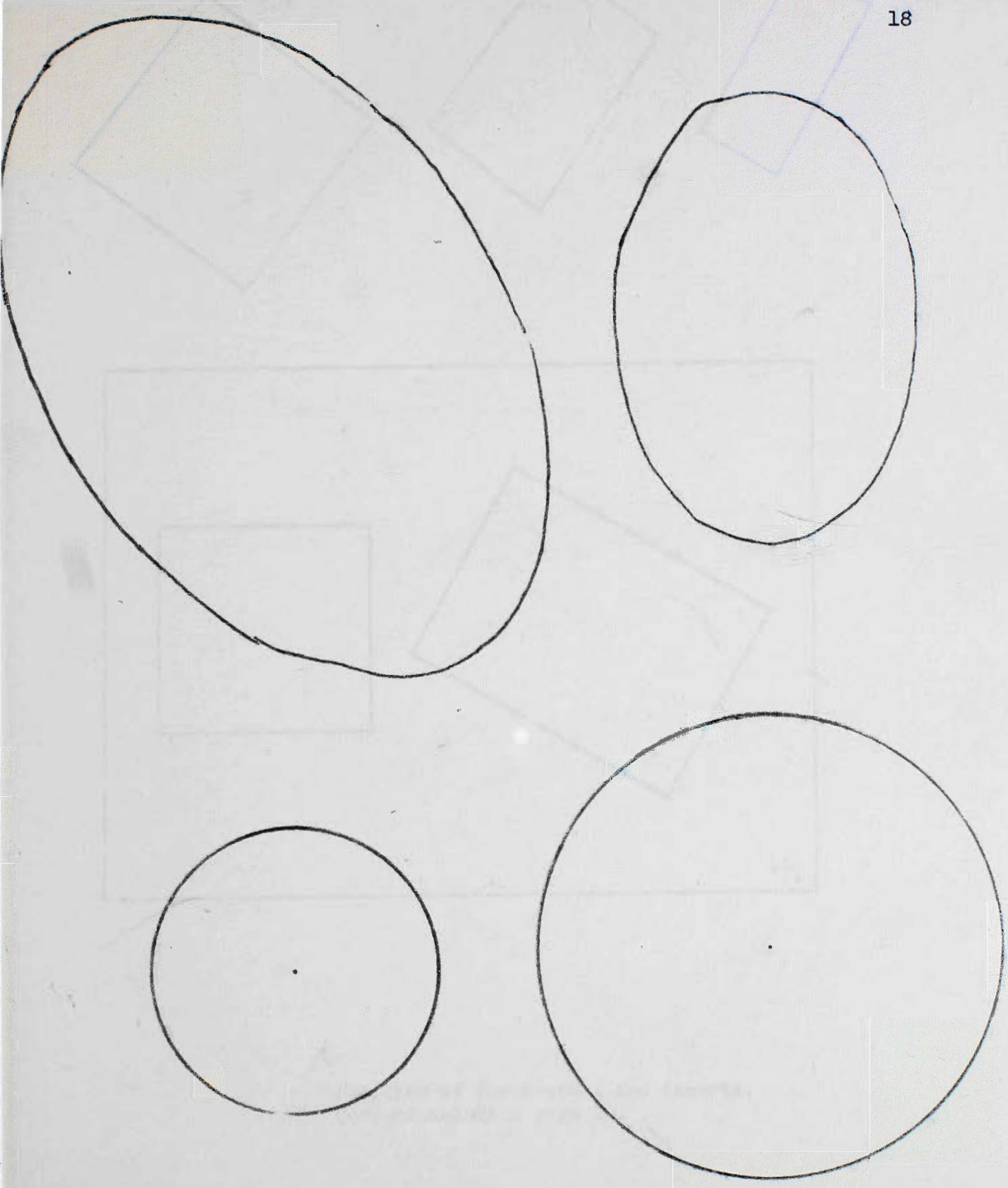
3. Size-Shape: Four forms, two circles and two lozenges of such size and shapes as will make their differences more obscure. In this board there is simple differences in size and minor differences in shape. The four inserts are painted white. The large lozenge and the small circle are at the bottom and the small lozenge and the large circle at the top, the lozenges on the left side and the circles on the right. The long diameter of the large lozenge runs from the lower left-hand corner towards the center but the long diameter of the small lozenge is parallel with the top and bottom edges of the board. The long diameter of the large lozenge is 5 1/8", the width is 3¼", and the long diameter of the small lozenge is 3 3/16" and the width is 2 1/16".

4. Size Combination: A simple problem of combination of inserts to fill a recess. The three inserts are painted white. Each piece is 3 3/8" long, the small piece is 1 5/8" wide, the second piece is 2¼" wide, and the large piece is 2 7/8" wide. The sides of the smaller recess are parallel with the edges of the form board but the larger rectangular recess is not parallel with the edges of the board. The largest insert fills the smaller recess and the other two pieces

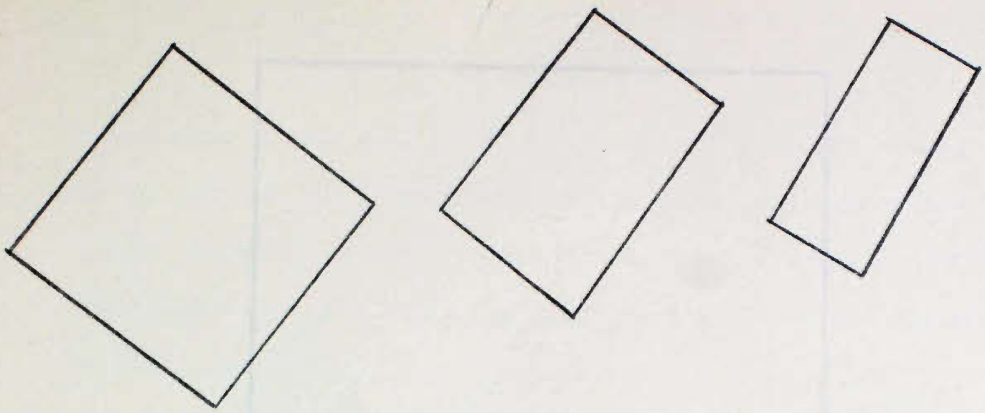


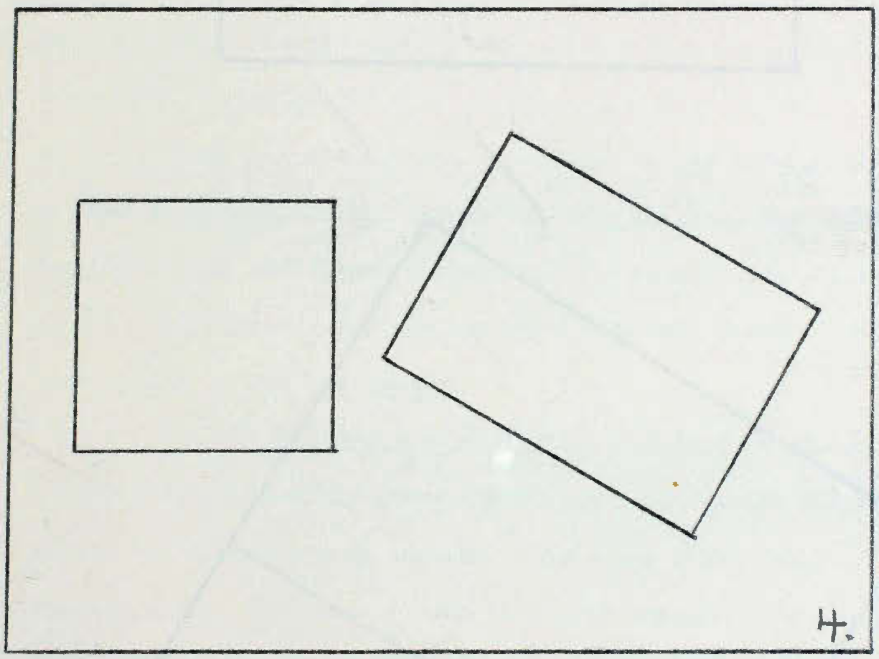


Position of Formboard 3 and inserts.
Correct solution page 18.

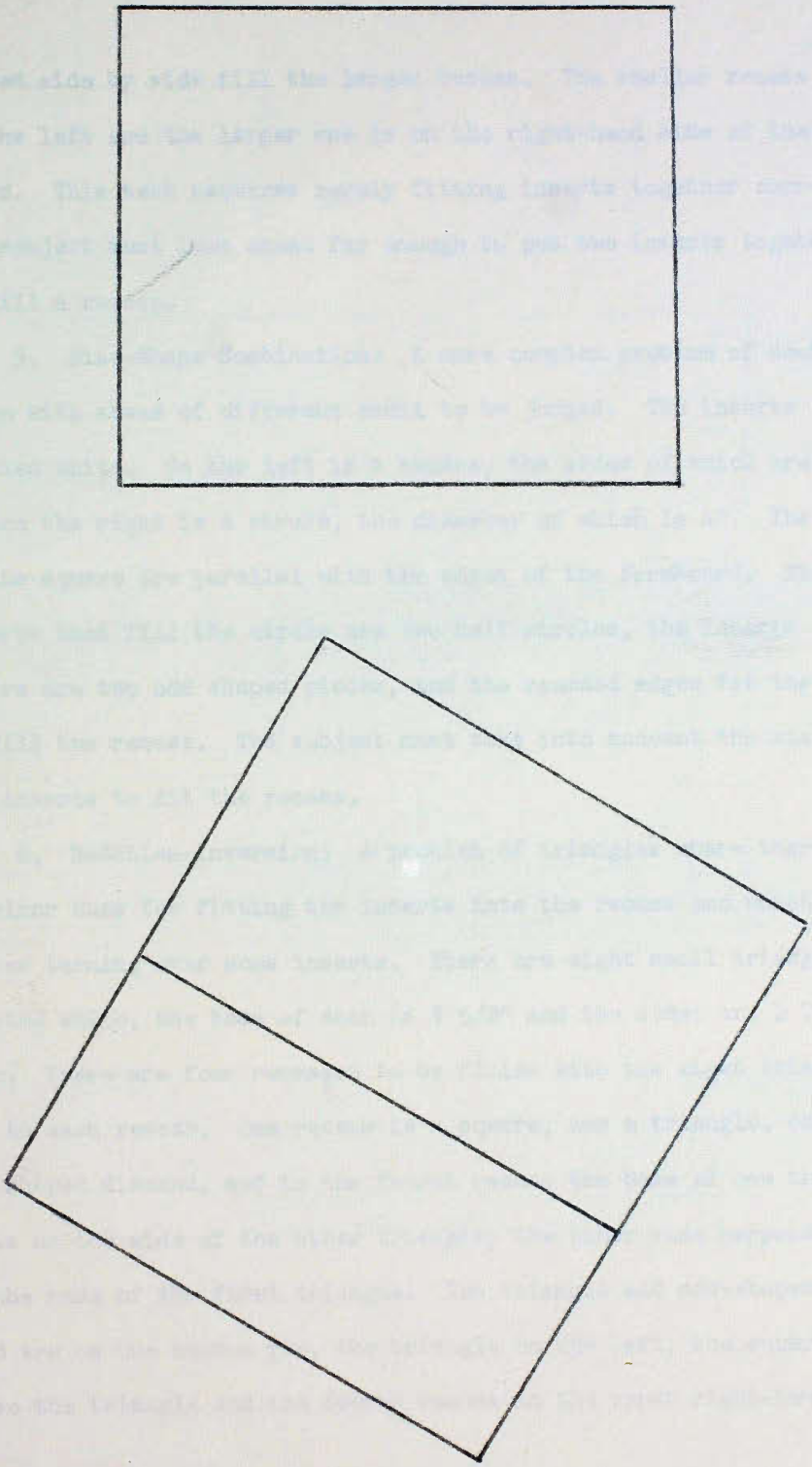


Correct Solution





Position of Formboard 4 and inserts.
Correct solution page 20.

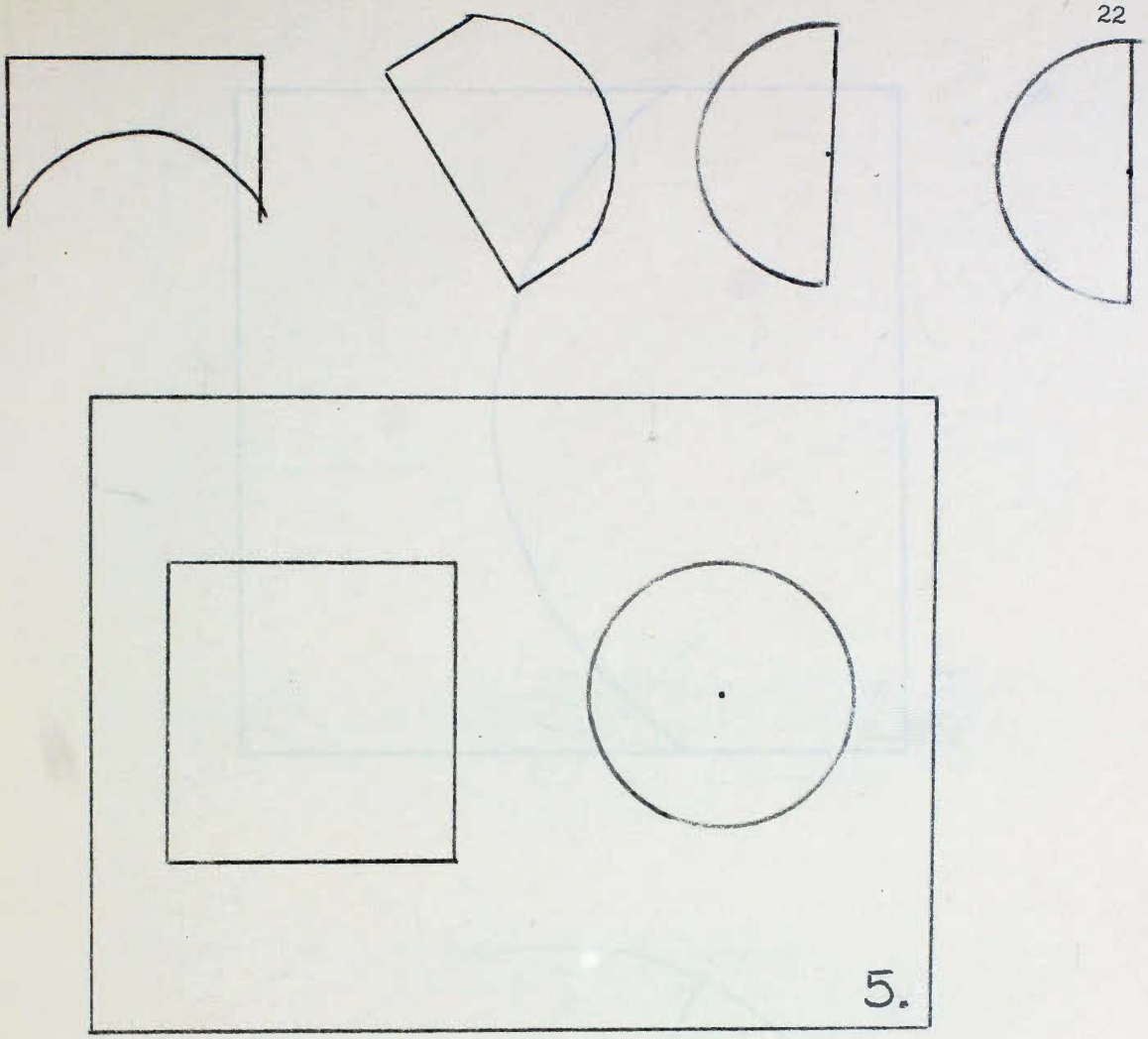


Correct Solution

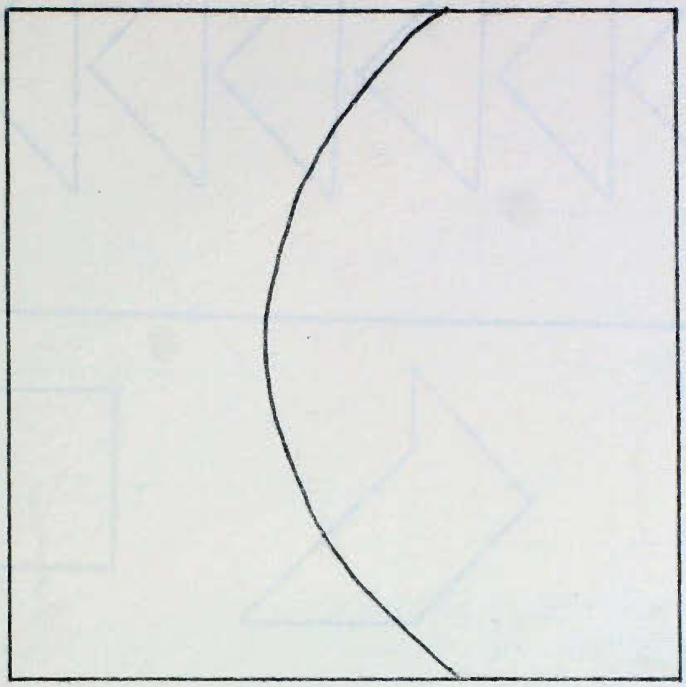
placed side by side fill the larger recess. The smaller recess is on the left and the larger one is on the right-hand side of the form board. This test requires merely fitting inserts together correctly. The subject must look ahead far enough to put two inserts together to fill a recess.

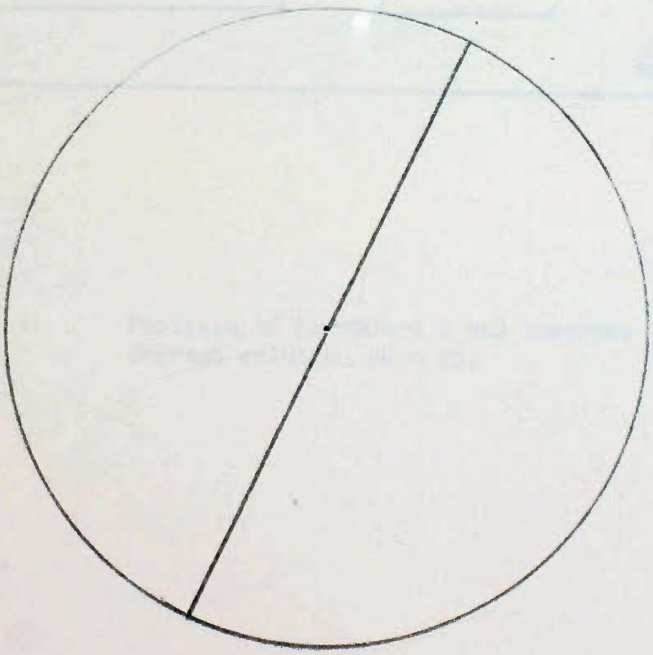
5. Size-Shape Combination: A more complex problem of combination with areas of different radii to be judged. The inserts are painted white. On the left is a square, the sides of which are 4" and on the right is a circle, the diameter of which is 4". The sides of the square are parallel with the edges of the formboard. The inserts that fill the circle are two half circles, the inserts of the square are two odd shaped pieces, and the rounded edges fit together to fill the recess. The subject must take into account the shape of the inserts to fit the recess.

6. Rotation-Inversion: A problem of triangles where there are no minor cues for fitting the inserts into the recess and which involves turning over some inserts. There are eight small triangles painted white, the base of each is 2 5/8" and the sides are 1 7/8" each. There are four recesses to be filled with the eight triangles, two to each recess. One recess is a square, one a triangle, one an odd shaped diamond, and in the fourth recess the base of one triangle rests on the side of the other triangle, the other side perpendicular to the base of the first triangle. The triangle and odd-shaped diamond are on the bottom row, the triangle on the left, the square is above the triangle and the fourth recess in the upper right-hand corner,

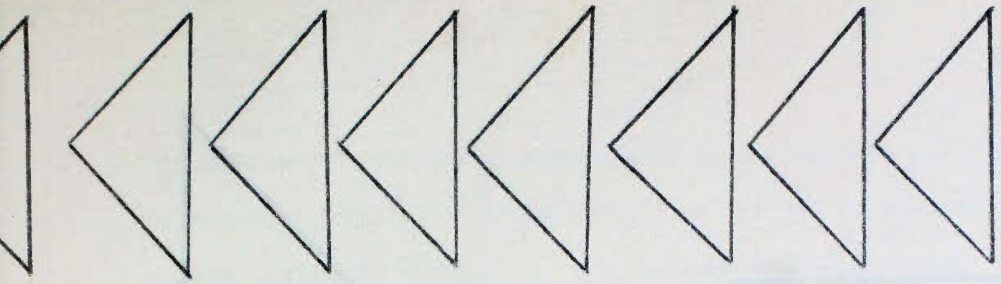


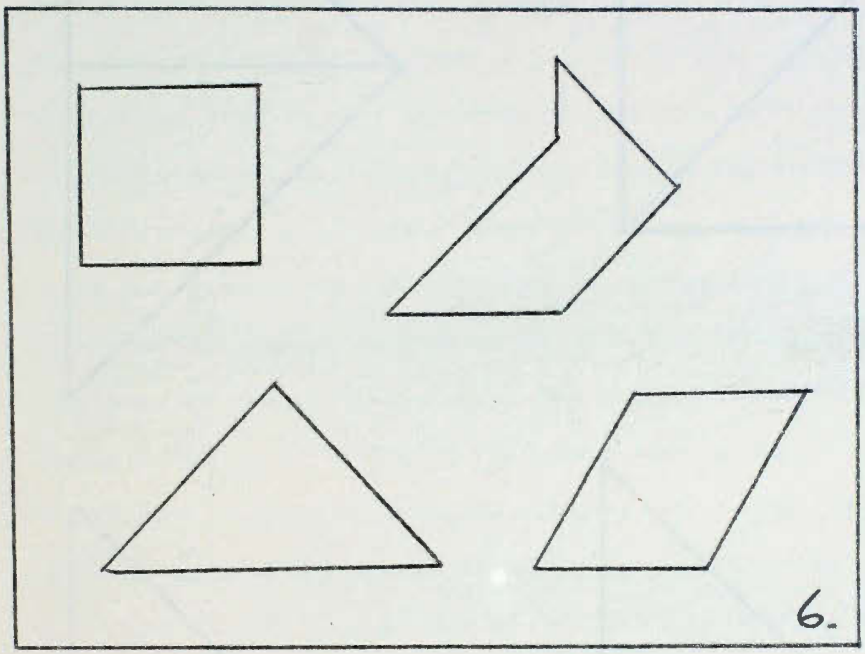
Position of Formboard 5 and inserts.
Correct solution page 23.



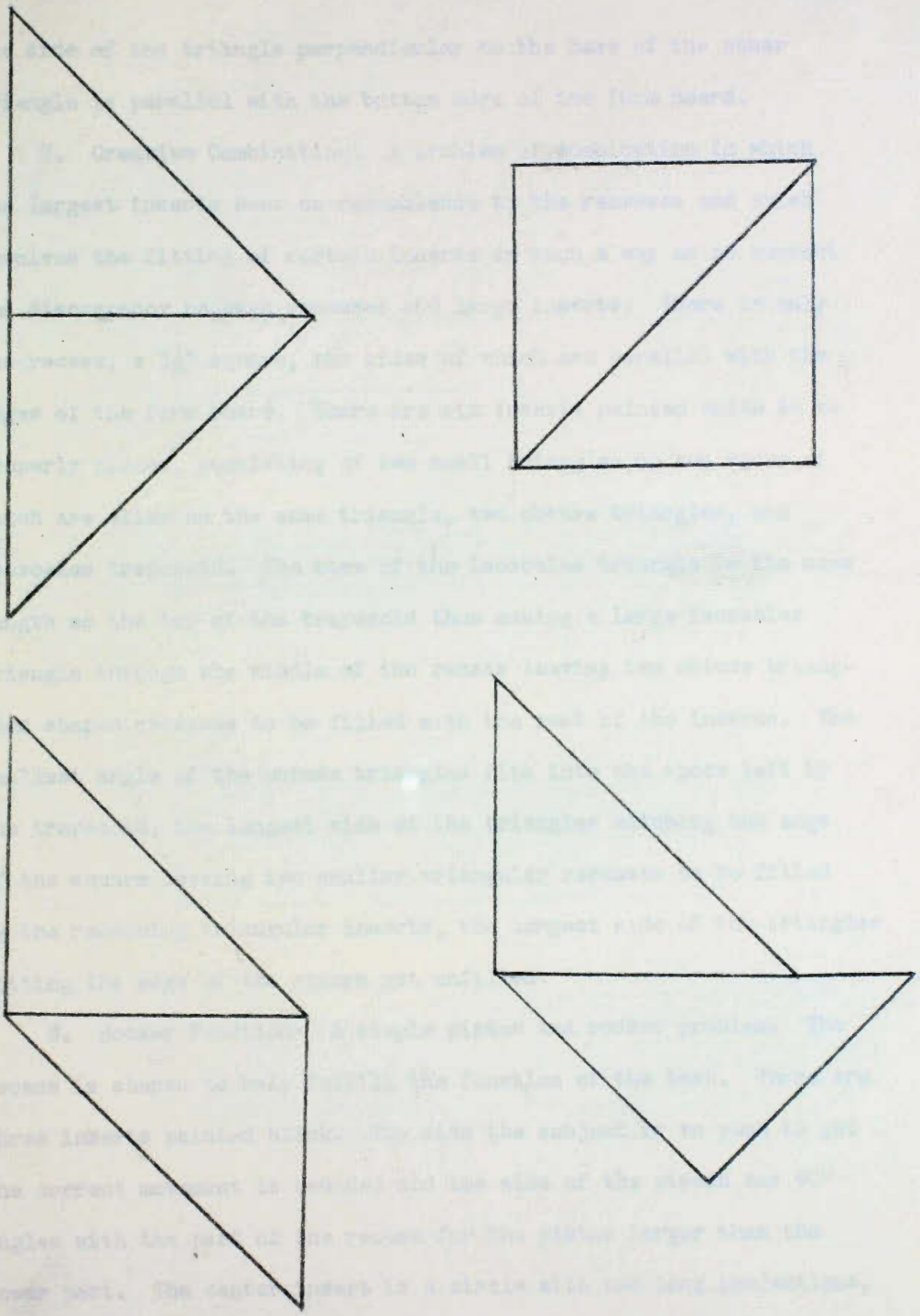


Correct Solution





Position of Formboard 6 and inserts.
Correct solution page 25.

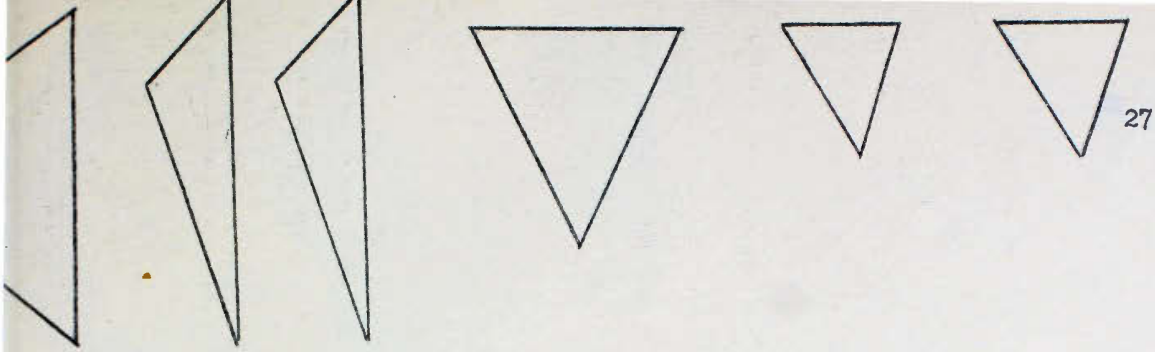


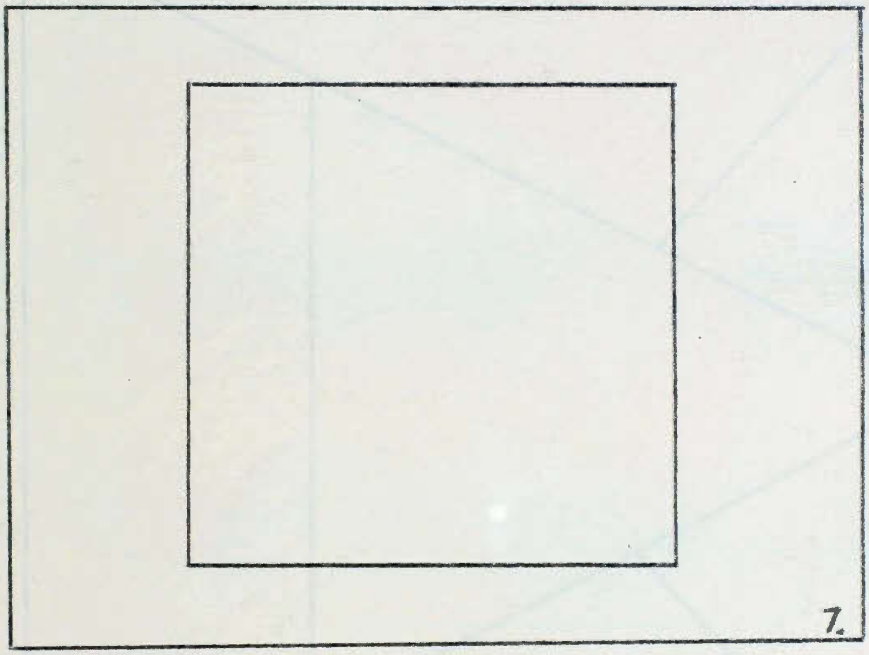
Correct solution.

the side of the triangle perpendicular to the base of the other triangle is parallel with the bottom edge of the form board.

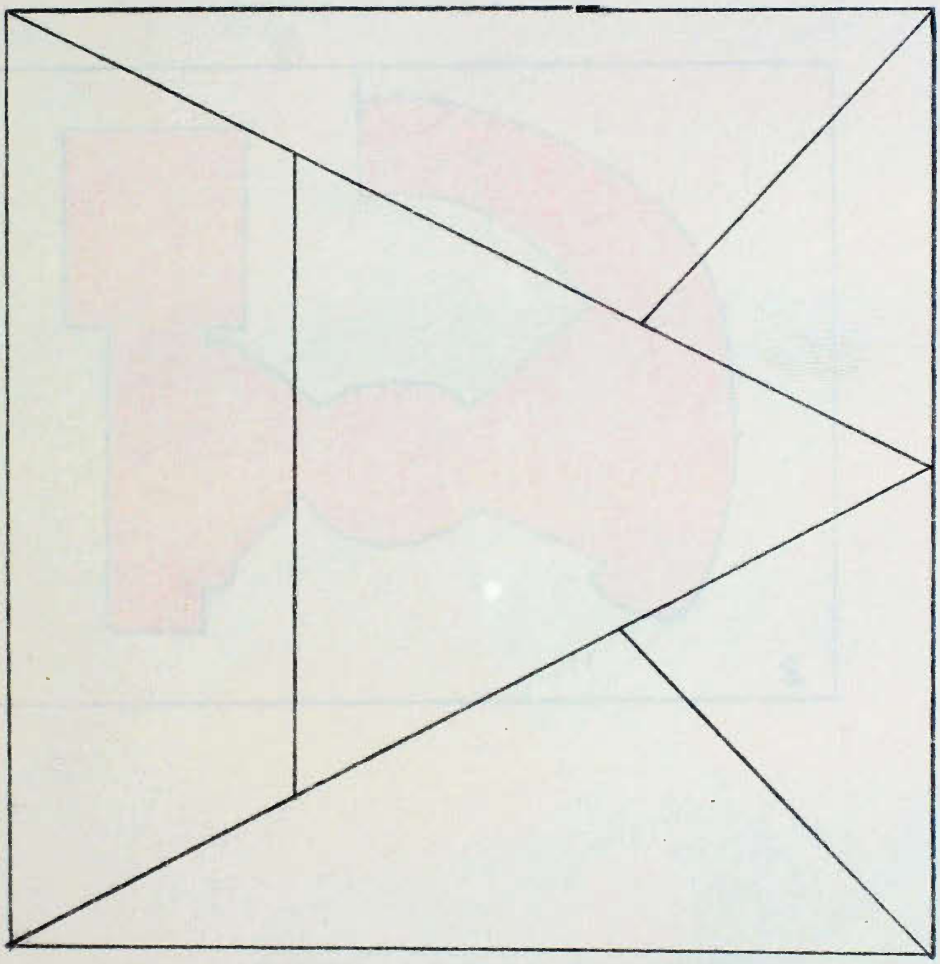
7. Creative Combination: A problem of combination in which the largest inserts bear no resemblance to the recesses and which involves the fitting of certain inserts in such a way as to correct the discrepancy between recesses and large inserts. There is only one recess, a $5\frac{1}{2}$" square, the sides of which are parallel with the edges of the form board. There are six inserts painted white to be properly placed, consisting of two small triangles no two sides of which are alike on the same triangle, two obtuse triangles, one isosceles trapezoid. The base of the isosceles triangle is the same length as the top of the trapezoid thus making a large isosceles triangle through the middle of the recess leaving two obtuse triangular shaped recesses to be filled with the rest of the inserts. The smallest angle of the obtuse triangles fits into the space left by the trapezoid, the longest side of the triangles matching the edge of the square leaving two smaller triangular recesses to be filled by the remaining triangular inserts, the longest side of the triangles fitting the edge of the square yet unfilled.

8. Rocker Function: A simple piston and rocker problem. The recess is shaped to help fulfill the function of the test. There are three inserts painted black. The side the subject is to push to get the correct movement is rounded and the side of the piston has 90° angles with the part of the recess for the piston larger than the lower part. The center insert is a circle with two long projections,

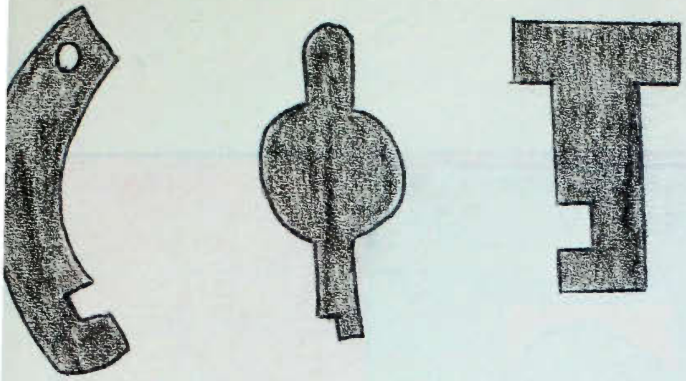


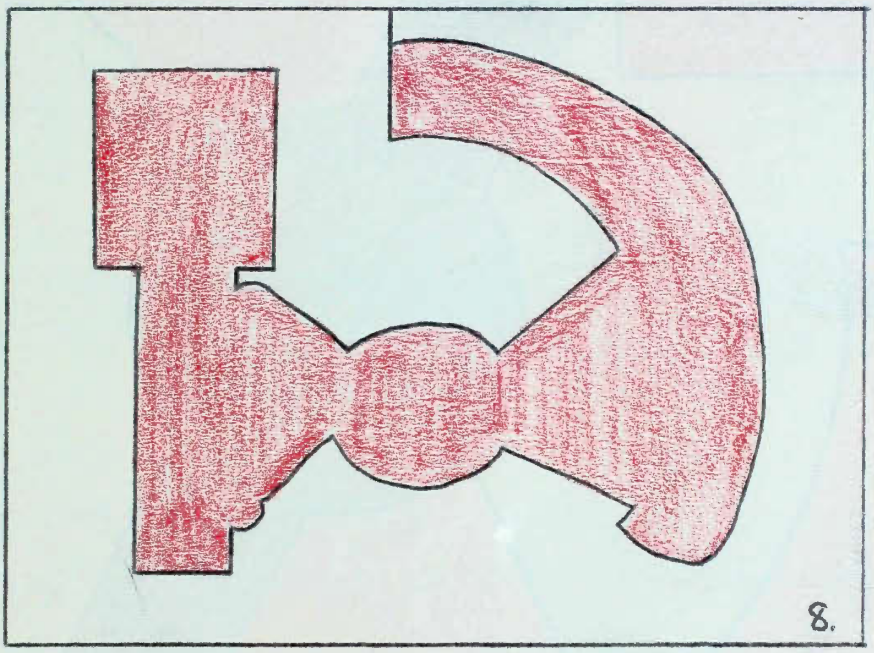


Position of Formboard 7 and inserts.
Correct solution page 28.

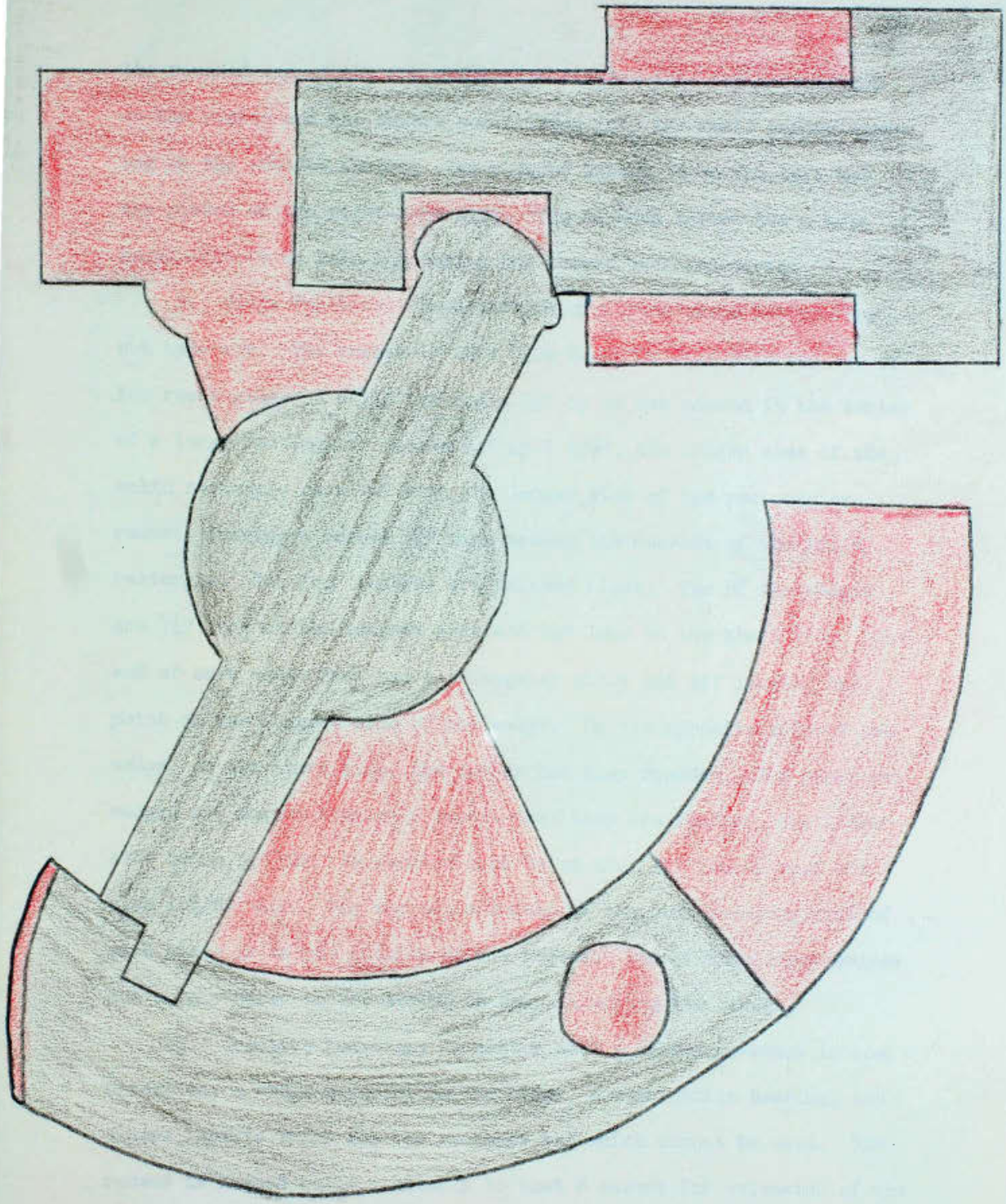


Correct Solution





Position of Formboard 8 and inserts.
Correct solution page 30.

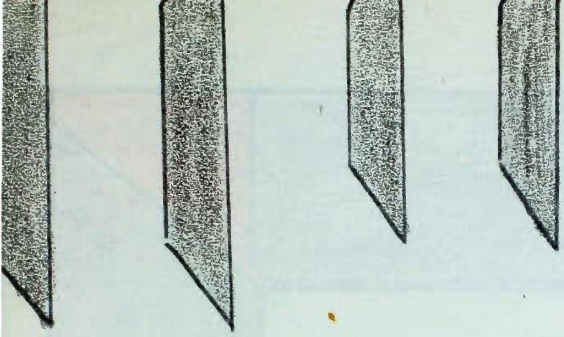


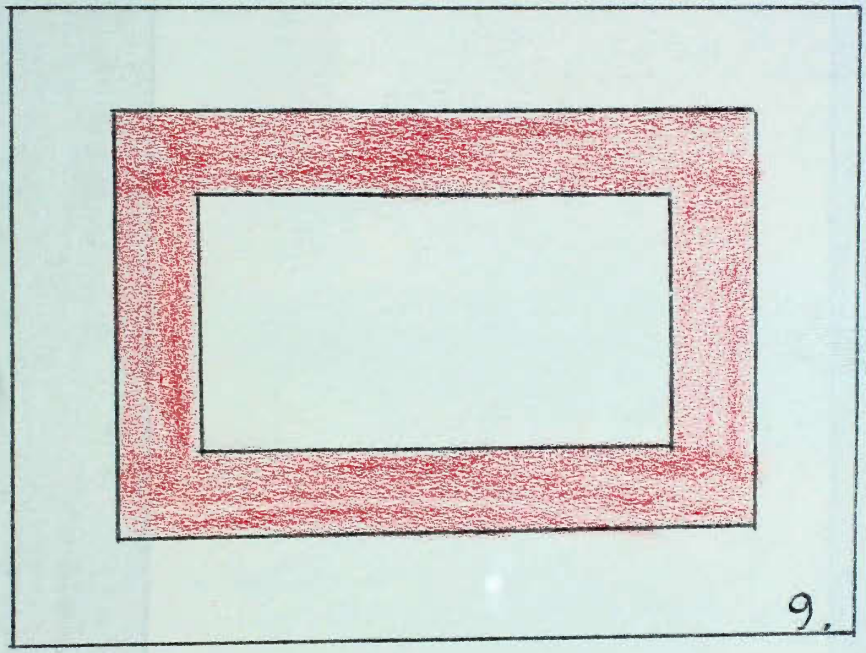
Correct Solution

the rounded end of the one projection fitting into the indentation on the piston and the square end fitting into the small square opening on the rounded insert. The rounded insert is on the left and the piston on the right-hand side. The rounded insert has a hole bored into it to help in pushing the insert back and forth.

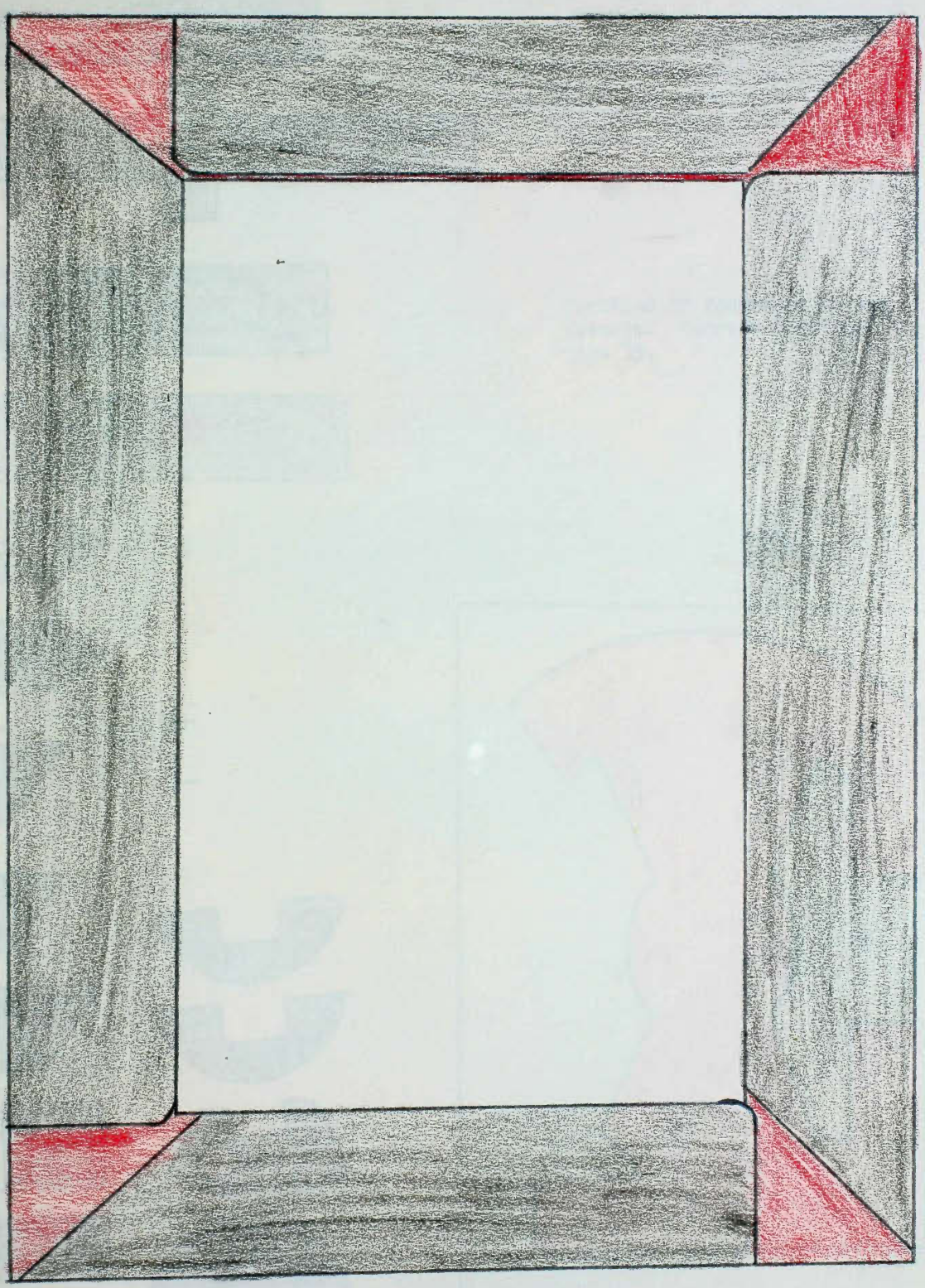
9. Wedge Function: Exploitation of forces where motions are not apparent. The recess of this form board is made with the aid of two rectangles. A solid rectangle $6\frac{1}{2}$" by 4" was placed in the center of a large rectangular recess $9\frac{1}{4}$" by 6 5/8", the longer side of the solid rectangle parallel with the longer side of the rectangular recess, leaving a recess $1\frac{1}{4}$" wide around the outside of the solid rectangle. The four inserts are painted black. Two of the wedges are $7\frac{1}{4}$" long on the longest side and $6\frac{1}{2}$" long on the short side. One end of each wedge has had a triangular piece cut off leaving the point on the longest side of the wedge. On the opposite side of the wedge, on the short side, the corner has been rounded. The other two wedges are shaped similarly except that they are shorter, the longer side being $5\frac{1}{4}$" and the shorter side being 4". All four pieces are each 1 3/8" wide. For correct solution of the test the long side of each piece is to the outside of the recess. One of the longer wedges has a hole bored in the center to help in moving the pieces.

10. Creative Function: A piston with a rocker movement in one direction, a wedge movement in the other, a combination bearing, and excess inserts which fit the recesses but which cannot be used. The recess is shaped quite similarly to test 8 except for extension of the recess at the bottom and the piston faces downward instead of upward.



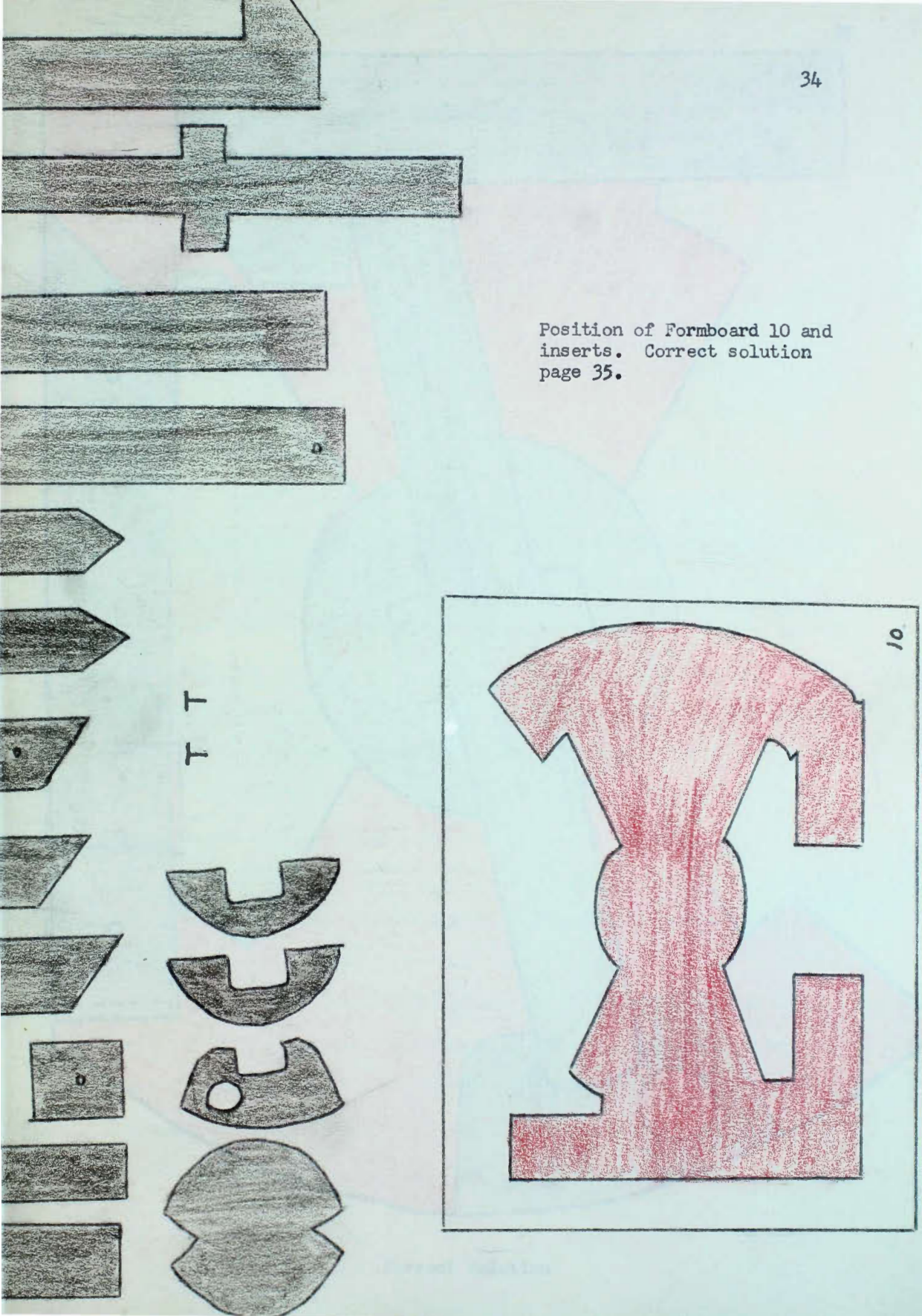


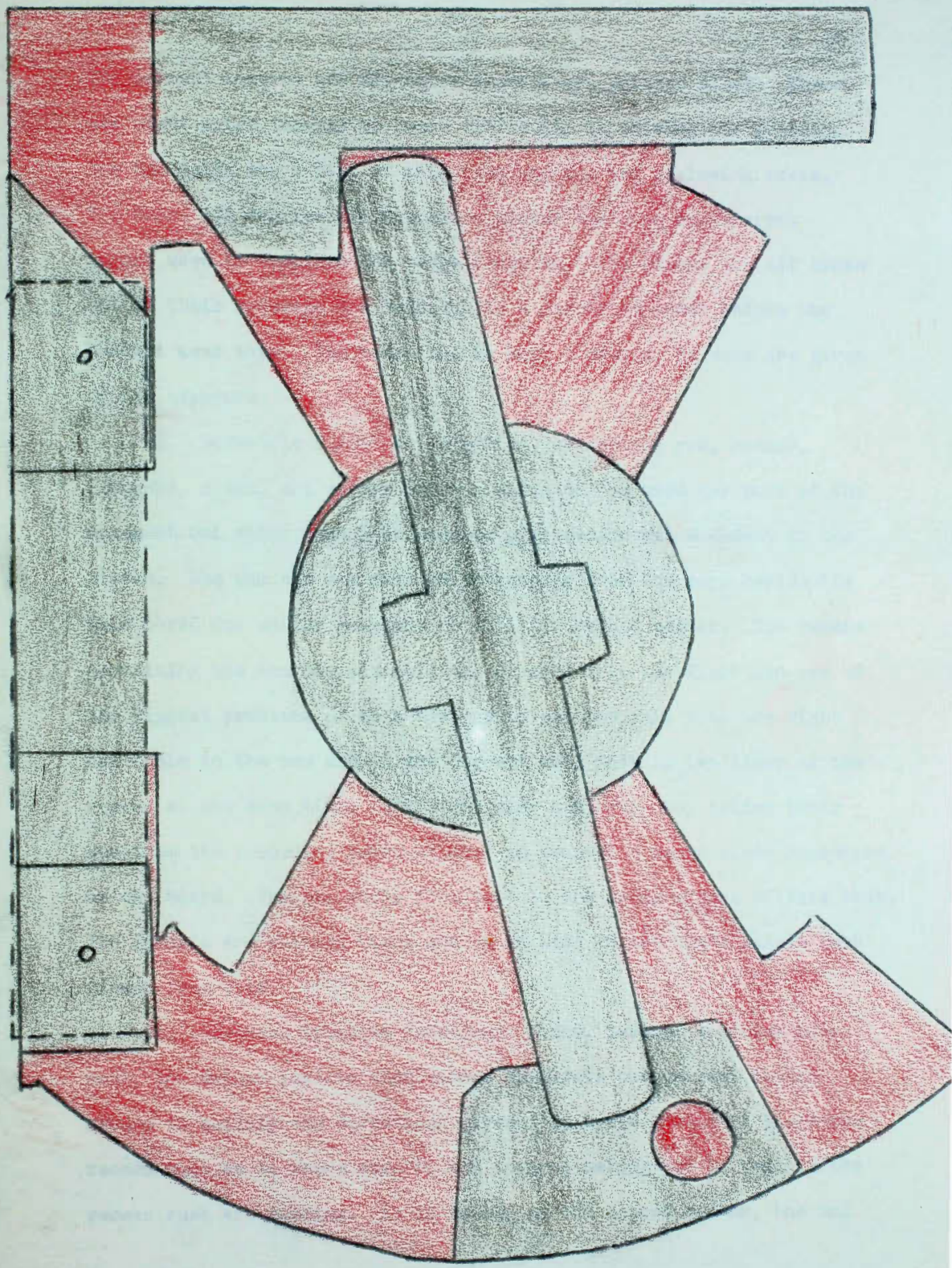
Position of Formboard 9 and inserts.
Correct solution page 33.



Correct Solution

Position of Formboard 10 and inserts. Correct solution page 35.



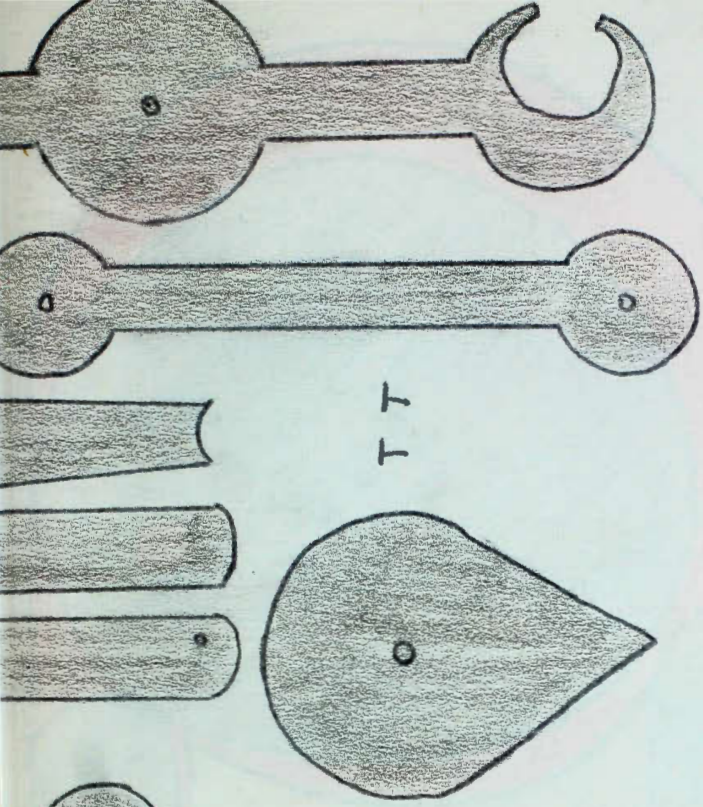


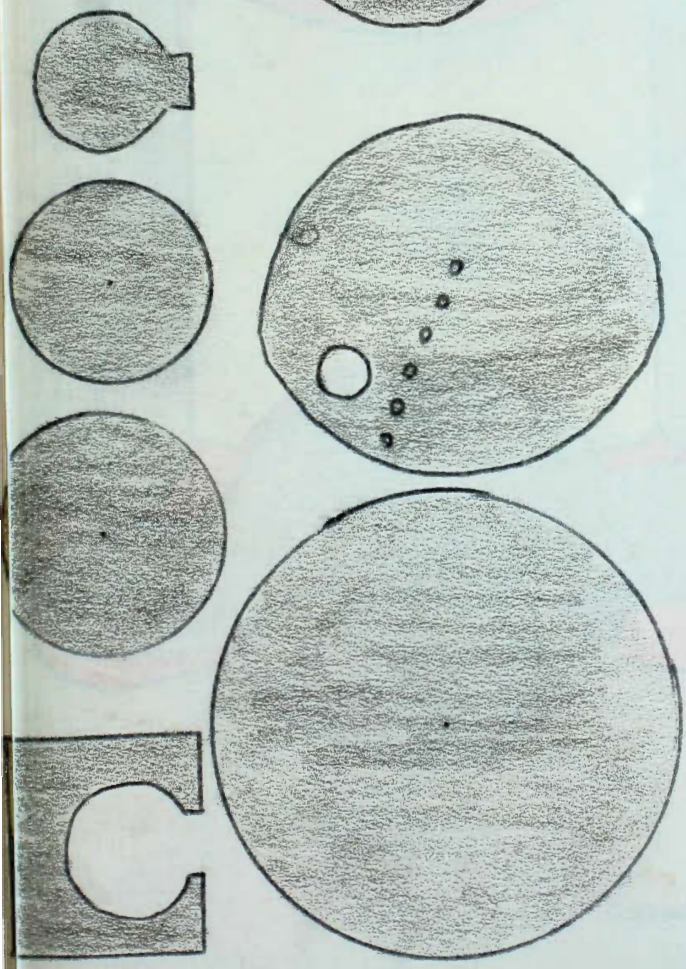
Correct Solution

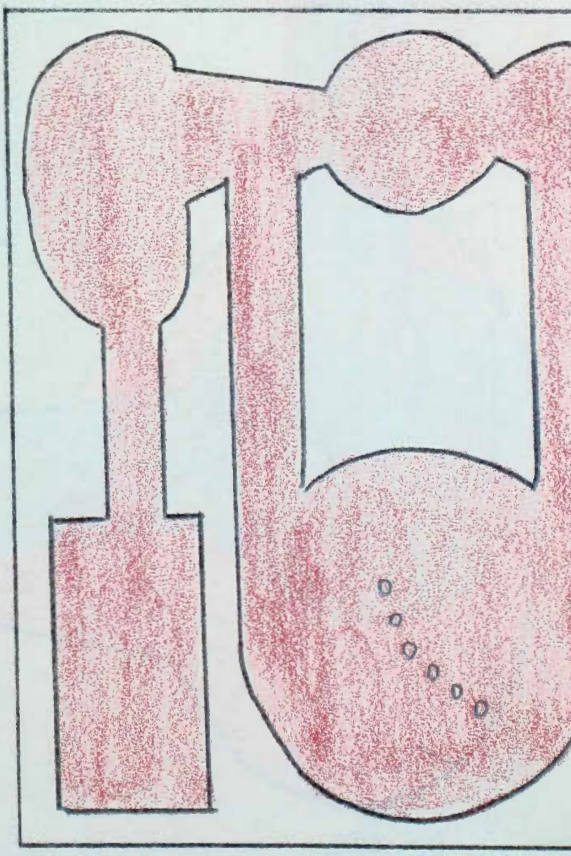
The correct inserts and all extra inserts are painted black. There are eight extra inserts to cause difficulty in solving the problem. The two nails are a part of this test and the two following tests. The board 10A carries all the extra pieces and also one correct insert used to complete the wedge function. The pieces are all taken out of their recesses and laid out in a specified order before the subject sees them. The board 10A is kept hidden so no cues are given by its exposure.

11. Eccentric Function: A piston, connecting rod, rocker, plungers, a cam, and excess inserts which can be used for part of the movement but which cannot be used to give reciprocal movement to the piston. The cam has six nail holes running from the edge beside the hole bored for easier movement to a little beyond center. The recess containing the cam has six nail holes bored into the floor and one of the biggest problems of this test is to get the nail into the right nail hole in the cam and in the correct nail hole in the floor of the recess at the same time. There are many subjects who, taking their cue from the preceding board attempt to connect a wedge piece somewhere on the board. The two nails from test 10 are again a part of this test. The inserts are painted black and as in test 10 the board 11A is kept hidden from view.

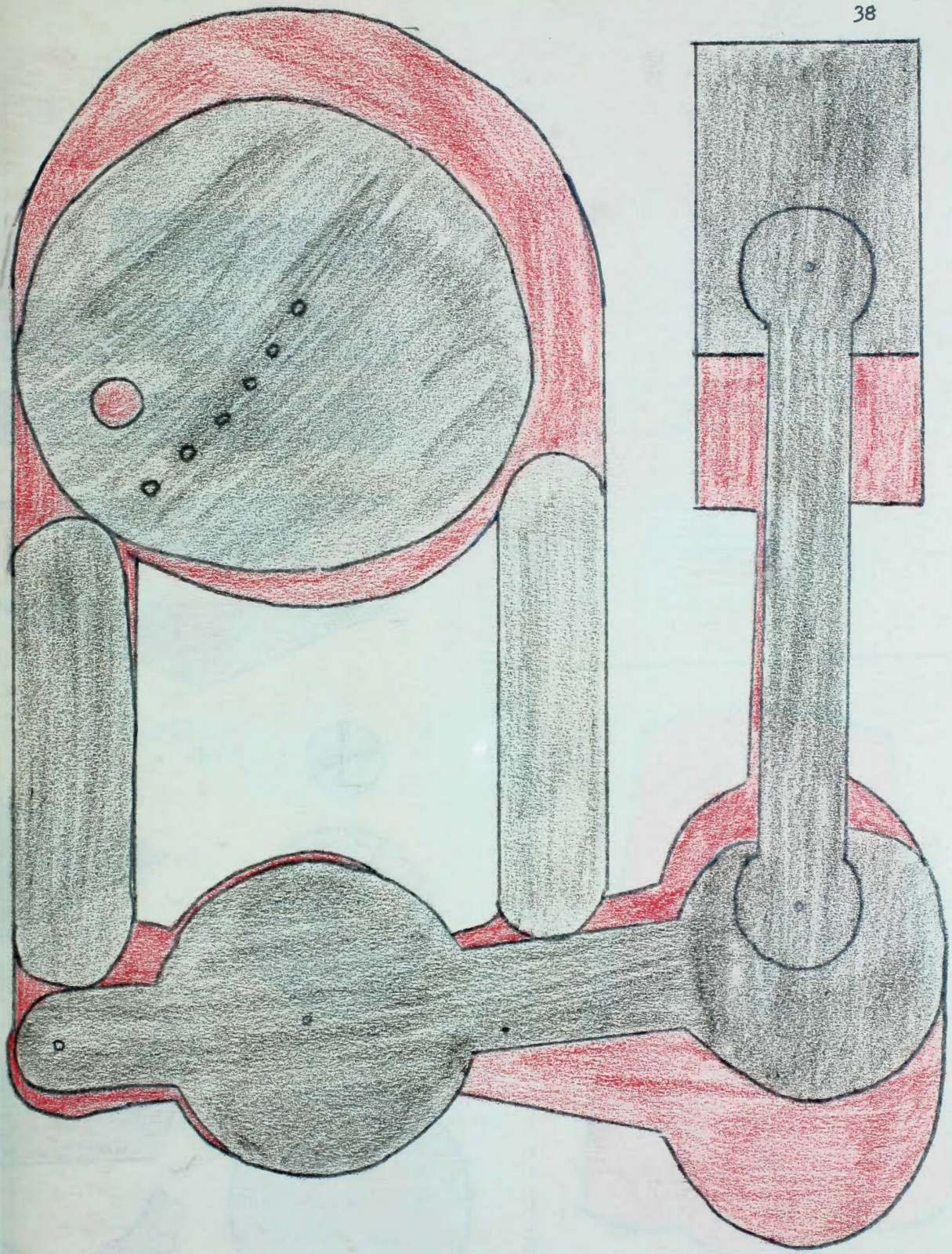
12. Advanced Creative Function: A cam, levers, a rocker movement, and excess inserts that cannot duplicate the correct action needed to satisfy the directions given. In tests 10 and 11 there are recess cues as to where some of the inserts belong but in test 12 the recess cues are missing. It is merely an odd shaped recess, the top



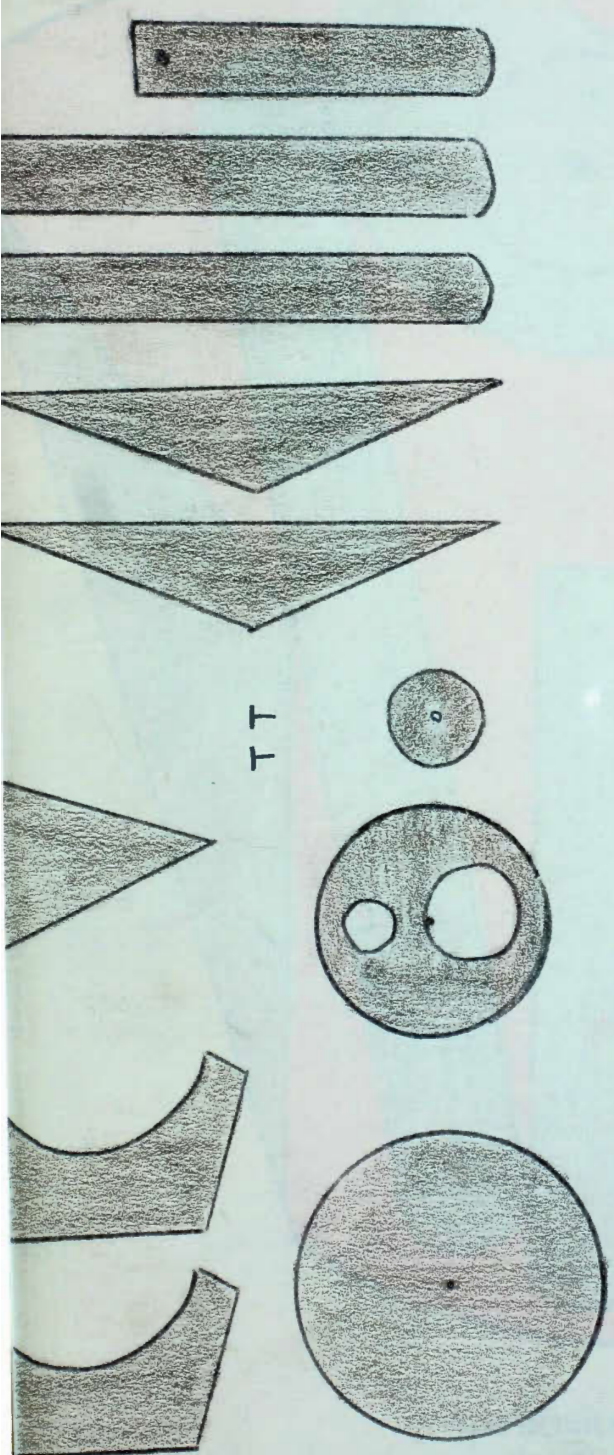




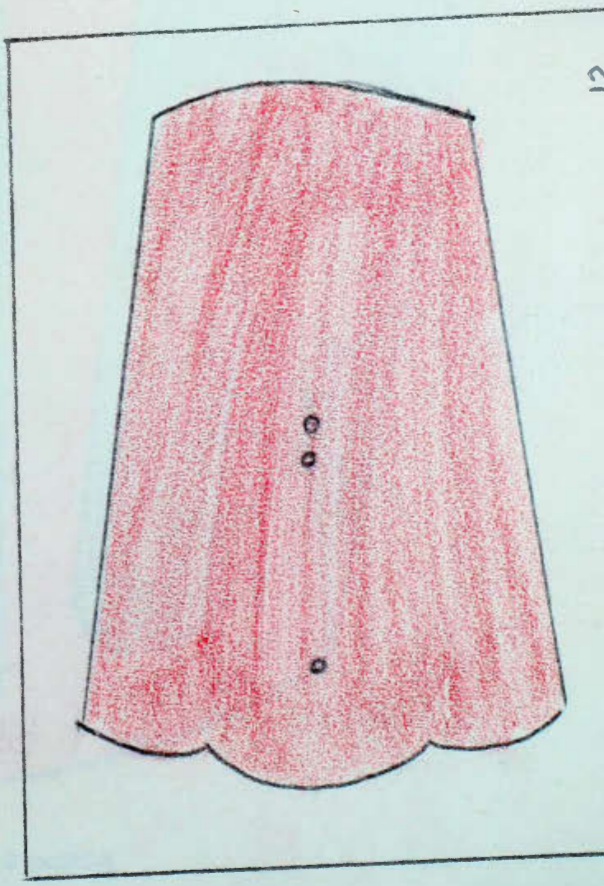
Position of Formboard 11 and inserts. Correct solution page 38.

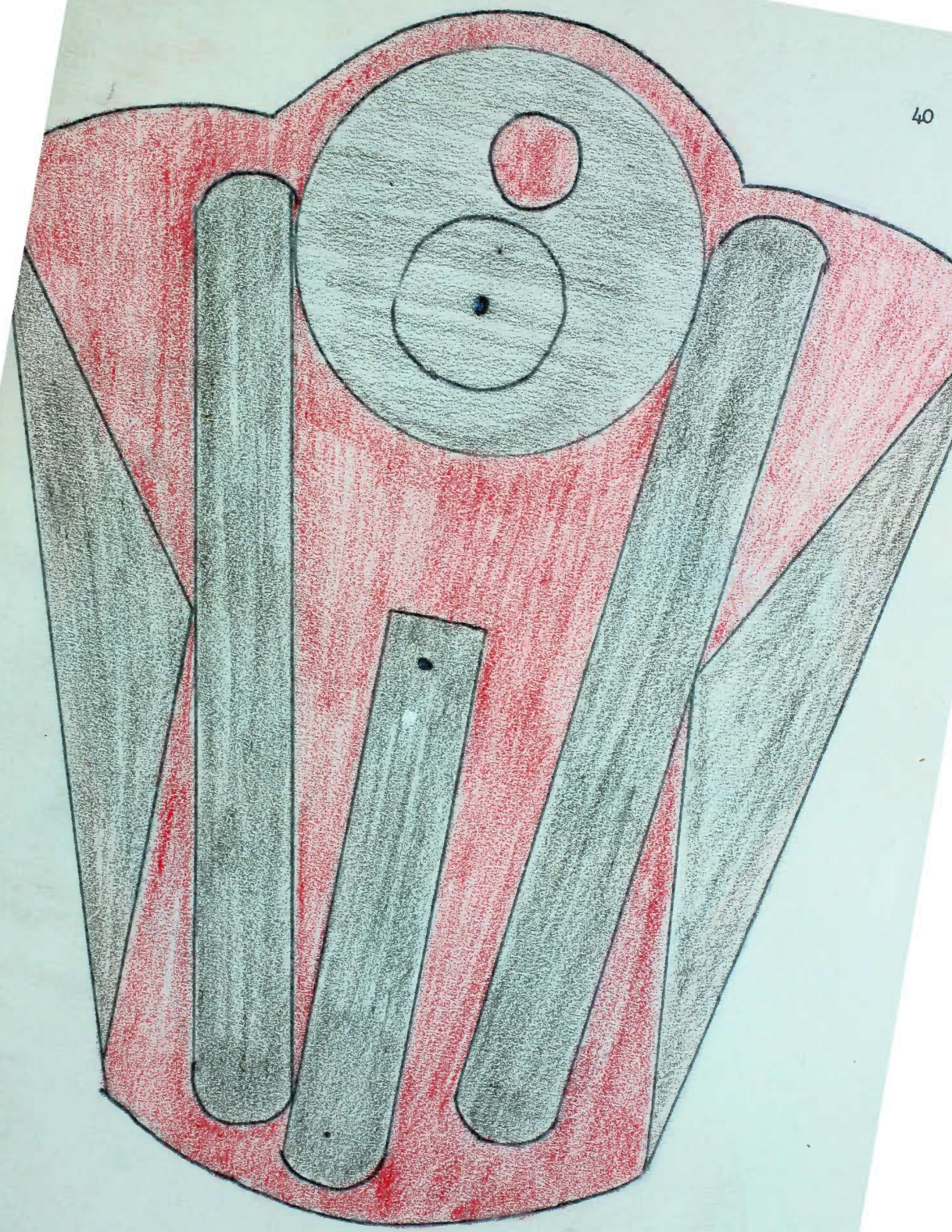


Correct Solution

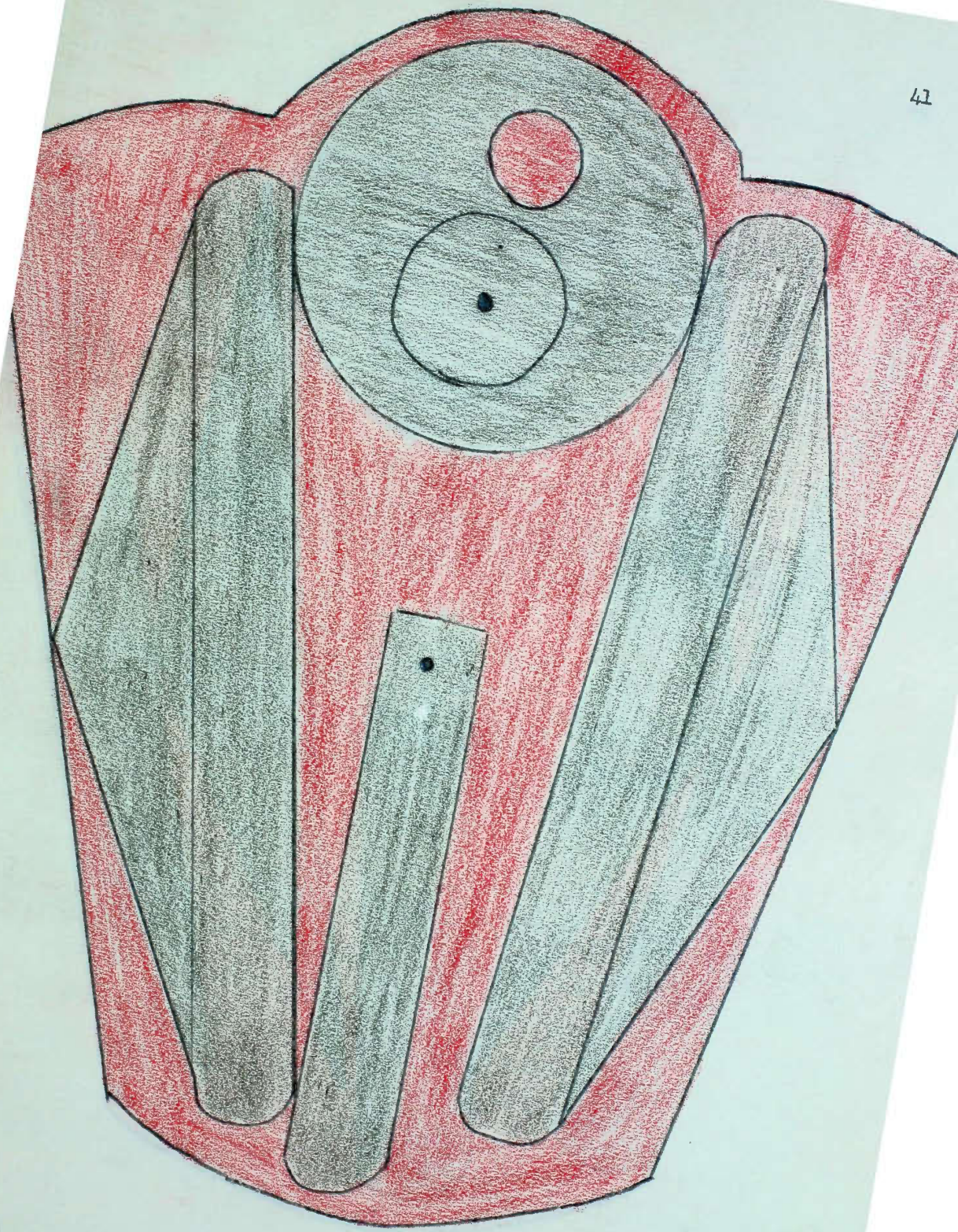


Position of Formboard 12 and inserts. Correct solutions pages 40, 41, and 42.

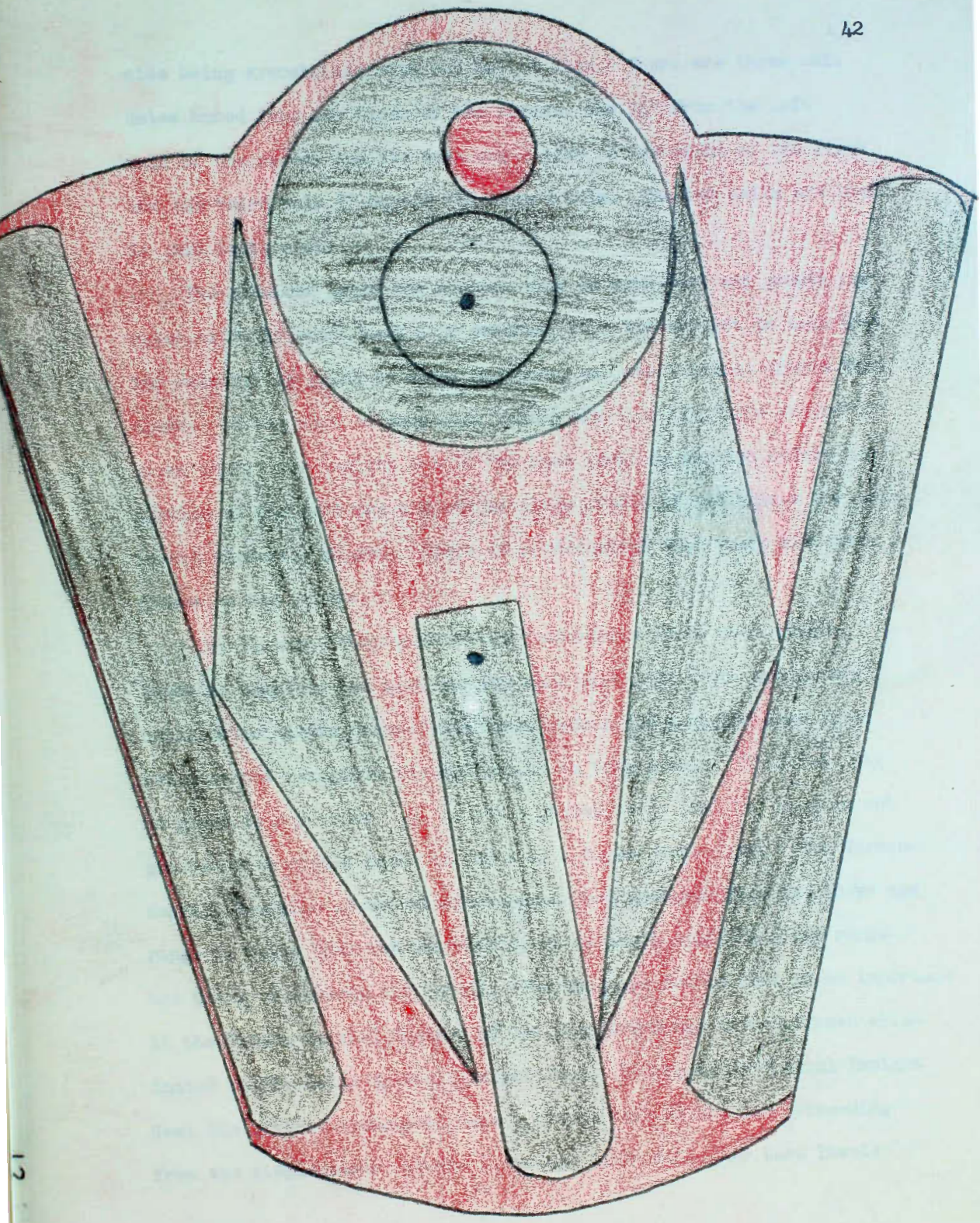




Best Solution 18 points



Good Solution 18 points



Spurious Solution 13 points

side being symmetrical with the bottom side. There are three nail holes bored into the floor of the recess, one $2\frac{1}{2}$" from the left wall of the recess and the next one 2 5/8" to the right of the first and the third hole $\frac{1}{4}$" beyond the second hole. The two nails are used in this test, also.

The examiner keeps the correct time on each test and counts the moves silently until test 9 is reached, when the subject is told that the examiner is going to count his moves and the examiner counts them aloud. If the test is completed before the time limit the subject is given credit for insight but if the time limit is reached before completion of the test the subject is given credit of two points for each insert correctly placed. There is no deduction when there are extra pieces in the recess.

The Pintner-Paterson Scale of Performance Tests used simple tasks to complete the test and the more difficult tests are merely complicated simple tasks. The Cornell-Coxe has a higher task level but does not recognize the difference in task levels. This test may be given in any order. It is not a graded test, that is, it does not proceed from simple tasks to tests of a higher task level. The Terman-Merrill Revision of the Stanford-Binet Test proceeds from the lower age range to tests that tax the ability of an adult. In the upper range the tests are nearly all verbal. The language factor that is so important in the Terman-Merrill Revision of the Stanford-Binet Test has been eliminated in the Kelly Spatial Insight Test. In the Kelly Spatial Insight Test the author attempts to make a graded test of ability proceeding from the simple tasks for the lower age limit to higher task levels

that require insight into the function before the correct solution can be found. The first seven tests are similar to the Pintner-Paterson Scale, proceeding from the simple task to complex simple tasks.

There is a change from Formboard 3 to Formboard 4 where the subject must make the adjustment from merely placing the correct blocks into the correct holes to placing two blocks into one recess in order to complete the task. Then, there is another great change from Formboard 7 to Formboard 8, where the subject must realize that not all the recess needs to be filled in order to complete the test. In the formboards with extra inserts there is a tendency for the subject to try to fill the whole recess instead of trying to find the correct function. The lower age subjects understand the idea of filling the recess with the blocks at hand but do not get the idea of "pushing something up and down over here to make something go up and down over there" without wanting to fill up all the space and then wondering why it doesn't work.

PROCEDURE:

Formulation of the record sheet: The first record sheets were very incomplete statements regarding the time and moves of the individual. The first idea was merely to count the number of moves and the time it took to complete the test, and if the test was not completed in the time allowed for it, the examiner showed the subject how to put it together and then the subject had a second trial in which the time and the number of moves were again recorded. The examiner, looking at the

record sheet later, saw that it did not give very much information concerning the individual's work. He had no idea how near right or how far wrong the subject was at the end of the time limit in either the first or second trial.

(System of scoring): This system of scoring proved very unsatisfactory so a staff conference consisting of the clinicians and Kelly was called and another record sheet was worked out. There were several traits that seemed important and necessary to be scored in order for the examiner to gain the most knowledge from his work. An insight score was isolated, that is, each insert correctly placed at the time limit, counted two points except on test 10 when three extra points were given for the wedge piece correctly placed. The nails counted the same as an insert giving two points for each mail correctly placed. The insight score on this sheet was taken on only the first trial. Time and moves were also isolated and a combined score of these two traits is called the efficiency score. This name was used because it most nearly showed the efficiency of the subject. If he could put the test together using only a few moves and only a portion of the time allotted to the test he was more efficient than the subject who made a number of useless moves and had to be stopped at the time limit. The time limits were already set up and the author saw no need for changing them. An arbitrary number of moves per minute was decided upon, one move every ten seconds. From observation this seemed to be about the rate a person worked who was not having particular difficulty with the board. One-tenth of the amount of time on the total trials for each

formboard plus the number of moves made, make up the efficiency score for that test. The total efficiency score on all boards completed or attempted plus the cumulative score of the highest "too easy" test and the cumulative score of the lowest "too hard" test make up the total efficiency score.

A complete record sheet was worked out giving space for the name of the subject, name of the clinician, the date of the test, birth date of the subject including month, day, and year of birth, chronological age of the subject, score on efficiency and score on insight. Then, there is space for the individual record of each formboard giving insight score on the first trial, time on each trial and the moves. Below this space is room for the total insight score (cumulative score of the highest "too easy" test plus the amount of insight on the tests completed or attempted, and also space for the total efficiency score. On the right hand side of the record sheet a temporary scale of insight and efficiency was made. It was not very satisfactory because it was possible for the scores made on the test to run off the end of the scale.

On the back of each record sheet the instructions for scoring and using the record sheet were printed. An explanation of the efficiency score and the insight score was printed on the back also for those who were not familiar with the score sheet.

In order to be fair to the student taking the test, in the staff conference it was decided to give him a chance to put the test together as he thinks it is right before grading it, so the following standard

statement was reached: "You have just thirty seconds left. Make the best solution you can and I will score it as it stands thirty seconds from now." The clinicians giving the test noticed it took the subject closer to sixty seconds to make his best solution so in a following conference the time was changed to one minute instead of the thirty seconds.

As stated above, the insight score was taken on the first trial only. It was taken for granted that the subject could put the test together the second time after completing the test for himself or being shown by the examiner. Then, cases were found that failed to put the test together the second time within the time limit. This had to be taken into account because the original insight score was too high for those who failed on the second trial so the number of pieces right was scored on each trial. This scoring of insight made the temporary scale on the right hand side of the record sheet obsolete and it was eliminated on the present record sheet. The cumulative insight score as worked out on the second record sheet was also eliminated and the new insight scoring system was added. In using this new system it is possible to tell at a glance how far the subject was able to go before failing to put a test together within the time limit. The insight score cumulates as the test progresses and when the limit is reached the final cumulative score is the total insight score.

((Alternative solutions to test 12): After setting up the scoring system other solutions were discovered to test 12 that satisfied all requirements and had to be considered. The best solution and its alternative which is almost as good as the first solution was scored

the settled way, two points for each block correctly placed making a total of 18 points and the two solutions that are not as satisfactory were given 13 points each. The spurious solutions satisfy all requirements but they are inferior to the best solution. The subject is given credit of 13 points for his solution and the time it took him to discover the solution is recorded. Then the subject is shown the best solution and is asked to put it together again. If he puts the best solution in the recess he is given 18 points on the second trial.

The subject had no knowledge of the solution of the task except the directions given by the examiner so the first trial provides all the learning for the subject. This brought up the problem of learning for the second trial so a minimum of time was given for the second trial. The second trial proved a good measure of the real understanding of the subject into the problem which he is trying to solve. The second recording of the number of inserts correctly placed is not a true insight, that is, the subject is no longer trying to solve a problem but trying to remember how he solved it the first time or trying to remember how the examiner put it together; but a score of learning that has taken place during the trial-and-error method of solving the problem the first time. He puts the test together as he remembers it and if at the short time limit he fails to complete the test it shows that in some way he has failed to master the test. He probably was not observing the way the examiner put the test together for him or failed to remember how he had placed the

inserts to make the test function. Test number 9 seems to be one that causes trouble on the second trial more so than tests 10-12. The odd-shaped wedge piece on test 10 causes some trouble. The subject fails to observe just how it was placed in the recess and as it is nearly covered by the cross-piece, the function of it does not show up very plainly. In test 11, the nail causes a lot of trouble. The correct nail hole does not appear to be the logical one because most of the subjects are more interested in putting the nail through the hole that is closest to the center of the cam than in making an off-center function to complete the rest of the movement.

The test was given to 159 subjects between the ages of 9 months and 26 years. Seventy-six of the subjects were college students between the ages of 18-26 years, and eighty-three subjects were from 9 months to 18 years. The child 9 months old was unable to grasp the idea of the test. The command to "put the blocks into the holes" meant nothing to him. To him the blocks were objects to be conveyed to the mouth as quickly as possible. A reliable test was given to a child 14 months old. Sixty-eight of the tests were given in a small community forty-five miles south of Hays, Kansas. The small country town has a population of about 225. The major occupation of the surrounding country is farming. Thirty-nine of the sixty-eight were children living on farms, twenty-seven lived in the small town of Alexander, and two were children from Wichita who were staying in Alexander for the summer. In the group of eighty-three children, three sets of twins were tested, one pair four years and five months of age, one pair six years and six months of age, and the oldest

pair eight years and eleven months of age. The other fifteen tests below college level were subjects living in Hays. Sixty-three of the sixty-eight tests were given in the home of the author. A table was set up in a bedroom that could be shut off from the rest of the home. This eliminated a large per cent of the noise and usual interruptions that is found. The room was kept well-lighted, if the day was cloudy the electric light hanging just to the left of the subject was kept lighted. On hot days an electric fan was kept in the room to prevent fatigue and sleepiness caused from the room being stuffy and too warm. For children to whom the author was a stranger and who were timid, up to the age of about 4 years, the mother or a sibling was allowed to remain in the room with the child. Most of the children came with the idea of "playing with blocks" and were willing to cooperate with the examiner. One boy, 12 years old, had been teased by his father and was afraid the examiner was going to hurt him but recovered his poise after reassurance by the examiner and enjoyed the test.

In 1938, an attempt was made to correlate the test scores with scores made on an intelligence test. The Henmon-Nelson Test of Mental Ability for grades 3-8, and 7-12, and for college students was used for these age groups. For the primary grades, the Detroit Primary, Detroit Beginning First Grade, and the Detroit Advanced First Grade, tests of mental ability were used. The low correlation found shows the Kelly Spatial Insight Test is not measuring what the usual paper and pencil intelligence test measures. The tests were made on

50 subjects between the ages of 5 years and college age. Twenty-five of the tests were given by the author for a project in Genetics and Applied Psychology class, and twenty-five were given by another student in the same class.

Since there is no correlation between the test scores and scores on intelligence tests, the chronological age of the subjects was used on which to base the norms.

Preliminary statistical work on the material gathered showed the necessity for combining standard scores for both insight and efficiency to make one total score on the test. The mean and standard deviation of the insight scores and the mean and standard deviation of the efficiency scores were found using 50 tests on which to base the means and standard deviations. The individual insight score was subtracted from the mean of the insight scores and the difference divided by the standard deviation. The same was done for the efficiency score and the final scores were added together and divided in half to get a final single score which would be a combination of insight and efficiency. This procedure gave us a single score that was more informative of the performance of the individual.

This did not prove to be a very satisfactory method of combining the two scores so another system was devised that took into account the three divisions of the test. The division lines are drawn between Formboards 3 and 4, and between Formboards 7 and 8. The first group, consisting of Formboards 1, 2, and 3, has as its mental set placing inserts into the correct recesses, the inserts being

the same shape and but slightly smaller than the recess to allow easy completion of the task. The second group, consisting of Formboards 4, 5, 6, and 7, has a shift in mental set from placing blocks into their correct recess to placing more than one block into a recess to fill it. The last group, consisting of Formboards 8, 9, 10, 11, and 12, has a much harder shift of mental set--from completely filling a recess to placing blocks into a certain part of the recess to make a movement.

Because of the differences in ages and function of the task in each group, these divisions were arbitrary means of dividing the test to get the differences in standard deviations of the lower age limit, the middle age range, and the upper age limit. In order to combine insight and efficiency scores they must be in the same units. This could be done by dividing the test into the three divisions and then securing standard scores for both insight and efficiency on each division.

The mean of the lowest group for insight is 34 and the standard deviation is 5, and for efficiency the mean is 1342 and the standard deviation is 17. In the middle range the mean for insight is 134 and the standard deviation is 32; for efficiency the mean is 1160 and the standard deviation is 120. The mean of the upper range for insight is 258 and the standard deviation is 13, and for the efficiency scores, the mean is 587 with a standard deviation of 135.

The insight mean in the lower age limit was placed in a table and new scores one standard deviation on either side of the mean were added to the table. The insight means of the other two ranges

were also included in the table and new scores one standard deviation from the means were placed in the table. If the scores one standard deviation from the mean of two adjacent ranges permitted, second scores two standard deviations from the means were added. The intermediate scores were found by interpolation. The efficiency scores were found using the same method as for the insight scores. For the convenience of the examiner the insight and efficiency scores were placed in the same table (Table I). The means in each range were placed together on the table and the scores one standard deviation from the mean were placed in line with each other. Weighted scores were assigned to both the insight and efficiency scores, (Table I).

These weighted scores were used on each individual's test and were plotted on graphs, one for insight (Fig. 1), one for efficiency (Fig. 2), and for the third graph (Fig. 3), the weighted insight and efficiency scores were added and then plotted on the graph. The vertical lines across the page are months and those along the left side of the page are the weighted scores.

The scores were plotted on a graph scaled very large, the weighted scores progressed by ones, the age in months by threes, and then by averaging in each column the first tentative broken-line graph was formed. Then to smooth the graph line, the scores on the broken-line were smoothed first by threes, secondly, by fives, which is the present broken-line.

On the combined weighted scores graph the heavy black line is the final smoothing in order to plot the norms for the different ages,

(Table II). From this norm table, the examiner can judge the mental age of the subject in relation to his work on the performance test.

VALIDITY:

There is no way to check the validity of the test because there is no other test on the market comparable to it.

RELIABILITY:

The reliability of the test was found by using the split-half method of correlation and was found to be r = .33. The Spearman-Brown Prophecy formula, which predicts the reliability of the whole test, raised the reliability from r = .33 to r = .50. The Spearman-Brown Prophecy formula is $r_{nn} = \frac{n\, r_{11}}{1 - (n-1)r_{11}}$ in which the r_{nn} represents the correlation between n forms of a test and n parallel forms and r_{11} is the reliability coefficient. Each half of the test used in finding the correlation contained only six formboards which made an extremely short test, and because of this it did not lend itself to the split-half correlation.

WEIGHT SCORE FOR INSIGHT AND EFFICIENCY

Insight	Efficiency	Weight	Insight	Efficiency	Weight
9	1427	1	142	1130	41
12	1416	2	146	1115	42
14	1410	3	150	1100	43
16	1404	4	154	1085	44
19	1393	5	158	1070	45
22	1382	6	162	1055	46
24	1376	7	166	1040	47
26	1370	8	170	1025	48
29	1359	9	174	1010	49
32	1348	10	178	995	50
34	1342	11	182	980	51
36	1336	12	186	965	52
39	1325	13	190	950	53
42	1314	14	194	935	54
44	1308	15	198	920	55
48	1306	16	202	912	56
51	1304	17	206	904	57
55	1302	18	210	896	58
58	1301	19	215	888	59
62	1299	20	219	880	60
65	1297	21	223	872	61
69	1296	22	228	864	62
73	1294	23	232	857	63
77	1292	24	236	823	64
80	1290	25	239	790	65
84	1288	26	242	756	66
87	1287	27	245	722	67
91	1285	28	248	688	68
95	1283	29	251	654	69
99	1281	30	254	620	70
102	1280	31	258	587	71
106	1265	32	261	554	72
110	1250	33	264	520	73
114	1235	34	268	486	74
118	1220	35	271	452	75
122	1205	36	274	418	76
126	1190	37	277	384	77
130	1175	38	280	350	78
134	1160	39	282	333	79
138	1145	40	284	317	80

TABLE I

MENTAL AGE NORMS

Years												
25	145											
24	144	144	144	144	144	144	144	144	144	144	144	144
23	144	144	144	144	144	144	144	144	144	144	144	144
22	144	144	144	144	144	144	144	144	144	144	144	144
21	144	144	144	144	144	144	144	144	144	144	144	144
20	143	143	143	143	143	144	144	144	144	144	144	144
19	142	142	142	142	142	142	142	142	143	143	143	143
18	140	141	141	141	141	141	141	141	141	141	141	141
17	139	139	139	139	139	140	140	140	140	140	140	140
16	136	136	136	136	137	137	137	137	138	138	138	138
15	132	133	133	133	134	134	134	135	135	135	135	135
14	128	128	129	129	130	130	130	131	131	131	132	132
13	122	122	123	123	124	124	125	125	126	126	127	127
12	116	116	117	117	118	118	119	119	120	120	121	121
11	108	109	109	110	111	111	112	113	113	114	114	115
10	101	101	102	103	103	104	105	105	106	106	107	108
9	93	94	95	95	96	96	97	98	98	99	100	100
8	86	86	87	88	88	89	89	90	91	91	92	93
7	79	79	80	80	81	81	82	83	83	84	85	85
6	71	71	72	73	73	74	75	75	76	77	77	78
5	63	64	64	65	66	66	67	68	68	69	70	70
4	53	54	55	56	57	58	59	59	60	61	61	62
3	43	44	45	46	47	47	48	49	50	51	51	52
2	31	32	33	34	35	36	37	38	39	40	41	42
1				22	23	24	25	26	27	28	29	30
	0	1	2	3	4	5	6	7	8	9	10	11
	Months											

TABLE II

ADMINISTRATIVE TECHNIQUE:

Terman and Merrill say (12) that in order to secure a valid result in the use of the Stanford-Binet scales, three requirements must be satisfied: (1) the standard procedures must be followed; (2) the child's best efforts must be enlisted by the establishment and maintenance of adequate rapport; and (3) the responses must be correctly scored. It can hardly be said that any one of the three is more important than the others, for all are absolutely essential. Unless the tests are given in strict accordance with the procedures by which they were standardized, the examiner can never be sure what his results mean. If he has failed to enlist the subject's best efforts, the only thing certain is that the resulting score will be too low in some unknown degree. Unless he has learned to score the responses according to the rules which have been laid down, his data will not be comparable to the norms.

These three requirements may well be applied to any test that is given and they are very necessary in administering the Kelly Spatial Insight Test.

(The Surroundings): The best testing room is one in which the child feels at ease and one to which he is accustomed. An unused schoolroom is good because the small distractions such as other children walking by the room or slamming doors and children shouting to one another are all noises to which he is accustomed and also the atmosphere of a school room is one in which the child has learned cooperation with others and to do his best on work that

is assigned to him. The test is merely another kind of school work to be accomplished. The room should be well lighted so the inserts are plainly seen and the room should be heated to a comfortable degree of temperature that inspires the greatest effort without fatiguing the subject. The table should have a light top so the black inserts show up to the best advantage but shouldn't be white or the white inserts will fail to show up to the best advantage. The subject is seated across the table from the examiner at a comfortable height from the table.

(Presence of others in the room): Other people in the room is one of the greatest distractions to a child taking a test. The other person cannot sit for an hour without moving and this distracts the attention of the child or if he fails to move then the child looks to see whether or not the individual is still there. For a child over four years or even younger if he is not a timid individual, the best rapport is gained from the child when he is alone with the examiner. He can give all his attention to what the examiner is presenting and what instructions are given to him without wondering what the third person thinks he can do with the test. For the younger children the presence of the mother is advisable until the child becomes adapted to his surroundings and then withdraw, and with very young children it is necessary for the mother or another relative to be present during the entire test.

(Duration of the test): The test at the very shortest takes thirty minutes for the very young child and the usual time for the youngest group is 45 minutes. The examiner should have his material

arranged in the correct order of presentation before starting the examination so there wont be any delay during the interview. For the best presentation the material may be left in the box until needed and after the child has completed that board it should be laid aside and covered to withdraw his attention from that board and focus it on the board next in order. For the older children the test takes an entire hour and even longer, and for the upper range it is usually 1½ hours before the test is completed. The college age subject can complete the upper tests in approximately one hour.

If the subject becomes fatigued a resting period should be inserted into the examination and the test finished after the child has had a chance to rest. A standard examination cannot be given if the child is tired or sleepy.

(Order of presentation): The boards are numbered from 1-12 and are designed to be presented in that order. The examination should begin with a form board low enough so the child will succeed in putting it together himself, and extend high enough so the child fails to get any insight score on the second or third presentation of the test. Tests should not be given that are obviously too hard or too easy for the subject.

The final group of tests were administered according to these instructions:

INSTRUCTIONS FOR ADMINISTRATION

1. Arrange the box so that form boards cannot be seen by subject until they are taken out of the box. Remove all form boards from

sight of subject as soon as they have been completed--except subjects above twelve years.

2. Do not let subject see inserts until they are out of their recesses. Keep formboard on edge and shielding inserts while arranging them.

3. Place formboards first with numeral in lower right corner for subject.

4. Rotate form board counter-clockwise between trials 90°.

5. Arrange inserts along top of board the same way for each trial, do not rotate with the board.

6. In scoring moves count as a move each direct comparison (with insert in or over a recess) as one move.

7. Record the number of moves in each trial.

8. For tests 1-6 inclusive say: "Put the blocks into the holes," and for test 7, "Put the blocks into the hole."

9. Arrange inserts from left to right according to the following criteria: First arrange according to diminishing size, if this does not clearly indicate order consider secondarily the diminishing simplicity of each figure. Arrange similar figures in similar position. Turn all pointed edges downward, (toward the subject).

10. Take notes on administration to aid in refinement of the test and correction of objectionable items.

11. In test 9 and subsequent tests say: "I am going to count aloud the number of moves you make."

12. In tests 8-12 say: "Put this together so that when you move something up and down here (move finger along arc in recess) it will make something go up and down here (move finger along piston track)."

13. In tests 9-12 add this to the other directions: "The nails are a part of the test and you may have pieces left over when you finish."

14. Administer tests 1-4 three times with rotations between trials. Administer all other tests twice.

15. Time limits for first trials:

Tests	1 - 4	2 minutes
Tests	5 - 6	4 minutes
Tests	7 - 8	8 minutes
Tests	9 -12	15 minutes

If the subject fails to complete test within the time limit, put it together for him and then let him have his other trials as follows:

Tests	1 - 8	1 minute
Tests	9 - 12	3 minutes

Administer all trials in tests 1-4.

16. One minute before the time limit on the first trial of each test say, "You have just one minute left. Make the best solution you can and I will score it as it stands one minute from now."

17. At the end of each trial score two points for each piece properly placed except test 10 where the outside cross piece counts five points.

18. In test 12 the two inferior solutions count 13 points.

19. Be sure to get age and exact date of birth of each subject. Keep exact record of all trials and scores. Be sure to get a standard test on each subject.

INTERPRETATION OF RESULTS

The standardization of the Kelly Spatial Insight Test gives

the Psychological Clinic another performance test that will prove useful in testing children brought to the Clinic. It will show the approximate mechanical ability of the child as he compares with the other children his age in western Kansas.

If the child does well on the test, completing most of the tests before the time limit and having some insight into the more difficult tasks, it means that that particular child will more likely succeed with mechanical types of work. Doing poorly on the test involves a number of influencing factors but with all environmental factors but with all environmental factors held constant he still has a poor score indicates a lack of mechanical ability.

If the child does well on this test but poorly on a reading type of test such as the Henmon-Nelson Test of Mental Ability, it may mean he has a verbal disability or has a poor background in verbal subjects.

This test can be used to find the mental age of a child in the Clinic as shown by his performance on this test when he has defective hearing. The instructions are few and very brief so a child with auditory deficiencies does not tire himself straining to hear a number of complex instructions.

It can be used to test children too young for a reading type of test. It does not require any reading ability, the only requirement is to understand simple instructions repeated by the examiner. The test is graded low enough to include children only 15 months old.

Because of the inability of the author to standardize the test

on children throughout the entire country the test cannot be taken as reliable for children of different sections of the country without a more extensive group of children being tested. There is no reason to believe, however, that the test is provincial. It may be necessary to exercise caution in applying the norms to other communities.

APPLICATION:

For the past year the test has been given to the students taking flight training on this campus. It has been used as one criterion for admittance to the Civilian Pilot Training Courses of the Civil Aeronautics Administration. The scores on the verbal intelligence test (Henmon-Nelson Test of Mental Ability for College Students, Forms A and B) were combined with the scores on the performance test (Kelly Spatial Insight Test) and were then correlated with the combined final flight rating scores. Matheny (7) found the correlation between the test scores and flight scores was $r = .405 \pm .077$. The cases which deviated markedly from the axis of positive correlation on the scattergram were analyzed clinically to discover what in each particular case had prevented the test score from accurately predicting the final flight score. On investigation it was discovered that some of the students who had not done particularly well in the tests turned out to be exceptionally good flyers. Others, it was discovered, had done more poorly in flight training than had been expected. On careful analysis it was dis-

covered some of these had exceptionally heavy scholastic and outside work loads.

These additional considerations were then embodied in the test score with such weights as seemed roughly to account for the incentives and obstacles discovered in the cases analyzed clinically. The formula finally used for all students was:

$$\text{Flying Aptitude} = \frac{V - \bar{V}}{\sigma V} + \frac{S - \bar{S}}{\sigma S} - \frac{W - 45}{10} \quad (+1) \text{ for career interest.}$$

Where: $\frac{V - \bar{V}}{\sigma V}$ = Verbal test deviation score in standard units.

$\frac{S - \bar{S}}{\sigma S}$ = Spatial Insight Test deviation score in standard units.

W = Weekly work load. (7)

When these modified test scores were correlated with the flight ratings, the correlation jumped to $r = .703 \pm .047$. This method was validated on one flight group. This is better than the prediction of scholastic success in college. In view of this experience, it should be perfectly safe to say that some considerable improvement could be made upon the Army's present wasteful elimination processes. If the new plan were as successful as in the experience of the work in the Clinic with the Civilian Pilot Training Program it should be possible to accept only 12,800 cadets from a group of 23,250 and come out with 10,000 trained pilots. At $3,800 each, the saving would be about $20,450,000 a year, or 65.7 per cent of the present total cost of "washouts" (7).

APPENDIX

Name_____ Age _____ Class_____ Date_____

I. FORM BOARD TEST

	Time	Moves
1	_____	_____
2	_____	_____
3	_____	_____
4	_____	_____
5	_____	_____

6. MOTOR SKILLS

Ball Bouncing	5		10	15	20
Asterognosis	1	2	3	4	5
Measurement	3	4	5	6	7
7. Sawing to line	4	8	16		
Finger tremor	xxxxXxxxxx				

8. Rocker Function:

Henmon Nelson for College Students

Percentile for Seniors. _____.

9. Static Function:

10. Creative Funct.

11. Eccentric Func.

12. Adv. Creat. Fun.

TOTAL INSIGHT SCORE (Add cumulative score of highest "too easy" test.

TOTAL EFFICIENCY SCORE (Note: Add cumulative score of highest "too easy" test--left hand figure -- and cumulative score of lowest "too hard" test--right hand figure--to total of scores made.)

Case _____ Clinician _____ Date _____

Date Birth	C. A.	Score	M. A.	I. Q.	Efficiency	Insight
						1300 0
I. Gross Shape:	Time: a2___ b1__ c1__	T___	1/10___			1280 2
4 (4)	Moves: a___ b___ c___	(6)	M___			1260 4
Insight___	Efficiency score				9	1240 6
						1220 8
2. Angle:	Time: a2__ b1__ c1__	T___	1/10___			1200 10
10 (6)	Moves a___ b___ c___	(9)	M___			1180 12
Insight___	Efficiency score				21____1332	1160 14
	Time:					1140 16
3. Size-Shape	Time: a2__ b1__ c1__	T___	1/10___			1120 18
18(8)	Moves: a___ b___ c___	(12)	M___			1100 20
Insight___	Efficiency score				36____1284	1080 222
						1060 24
4. Size Combin:	Time: a2__ b1__ c1__	T___	1/10___			1040 26
24(6)	Moves: a___ b___ c___	(9)	M___			1020 228
Insight___	Efficiency score				48____1236	1000 30
						980 32
5. Sixe-Shape Comb:	Time: a4__ b1__	T___	1/10___			960 34
32(8)	Moves: a___ b___	(8)	M___			940 36
Insight___	Efficiency score				62____1188	920 38
						900 40
6. Rotation-Invers:	Time: a4__ b1__	T___	1/10___			880 42
48(16)	Moves: a___ b___	(16)	M___			860 44
Insight___	Efficiency score				84____1140	840 46
						820 48
7. Creative Combin:	Time: a8__ b1__	T___	1/10___			800 50
60(12)	Moves: a___ b___	(12)	M___			780 52
Insight___	Efficiency score				108____1080	760 54
						740 56
8. Rocker Function:	Time: a8__ b1__	T___	1/10___			720 58
36 (36)	Moves: a___ b___	(6)	M___			700 60
Insight___	Efficiency score				126____ 972	680 62
						660 64
9. Wedge Function:	Time: a15_ b3__	T___	1/10___			640 66
74 (48)	Moves: a___ b___	(8)	M___			620 68
Insight___	Efficiency score				154____ 864	600 70
						580 72
10. Creative Funct:	Time: a15_ b3__	T___	1/10___			560 74
97 (23)	Moves: a___ b___	(20)	M___			540 76
Insight___	Efficiency score				194____ 648	520 78
						500 80
11. Eccentric Func:	Time: a15_ b3__	T___	1/10___			480 82
113 (16)	Moves: a___ b___	(14)	M___			460 94
Insight___	Efficiency score				228____ 432	440 886
						420 88
12. Adv. Creat. Fun:	Time: a15_ b3__	T___	1/10___			400 90
129 (16)	Moves: a___ b___	(18)	M___			380 92
Insight___	Efficiency score				____216	360 94
						340 96
						320 98
						300 100
						280 102
						260 104
						240 106
						220 108
						200 110

TOTAL INSIGHT SCORE (Add cumulative score of highest "too easy" test. [ ]

TOTAL EFFICIENCY SCORE (Note: Add cumulative score of highest "too easy" test--left hand figure -- and cumulative score of lowest "too hard" test--right hand figure--to total of scores made.) [ ]

Instructions for scoring and using score sheet

Most performance tests are comprised of tasks which require a comparatively low level of insight. Difficult tasks are devised merely by compounding tasks which an infant could accomplish if his attention span were great enough. This deficiency in performance testing has probably arisen out of the need for simple instructions which would circumvent verbal difficulties.

The present test is designed to sample the subject's ability to perform at various levels of insight. All of the tasks are in the field of spatial relations. A comprehensive measurement of a person's ability should also include tasks in other areas.

Each of the boards is named according to the type of insight required to solve it.

1. Gross Shape Discrimination
2. Angle--Discrimination of an attribute of shape
3. Size-Shape--Discrimination of space-filling fitness both in size and shape
4. Size Combination--Building areas to meet specifications
5. Size-Shape Combination--Building areas while disregarding minor shape cues.
6. Rotation-Inversion--Orientation
7. Creative Combination--Building areas without recess cues
8. Rocker Function--Devising an instrument to perform a functi
9. Wedge Function--Exploitation of forces where motions are no apparent
10. Creative Function--Use of wedge and rocker without recess c
11. Eccentric Function--Special form of the wedge and rotary mo
12. Advanced Creative Function--Eccentric and rocker without recess cues

Each of the first four tests is administered three times and each of the remaining tests is administered twice (See Instructions for Administration). Record time in seconds for each trial and ent on the "Time" line after "a", "b", and "c" respectively. Record th total time for all three (or two) trials after "T". Take one tenth (nearest whole number) of this total time and record it after "1/10 Record the number of moves in each trial and record on the "moves" line after "a", "b", and "c" respectively. Record the total number of moves on all three trials after "M". Add this score to the figur immediately above it (one tenth of the time in seconds) to obtain the "Efficiency Score". Enter "Efficiency Score" for the test.

Thirty seconds before the time limit on the first trial of eac test say, "You have just thirty seconds left. Make the best soluti you can and I will score it as it stands thirty seconds from now." At the end of the <u>first trial only</u> score two points for each piece properly placed except Test 10 where the outside cross piece counts five points and Test 11 where the pivot nail counts four points. I Test 12 the two inferior solutions count 13 points. This score is the "Insight" score. Enter it after "Insight".

Do not give tests which are obviously too hard or too easy. Cumulative scores are given on the blank for inferring optimum and minimum performances on all tests below or above a certain test.

Case _____ Clinician _____ Date _____

Birthdate _____ C.A. _____ Score _____ M. A. _____ I. Q. _____ Efficiency _____ Insight _____

1. Gross Shape: Time: a2___ b1___ c1___ T___ 1/10___
   4-8-12 (4) Moves: a___ b___ c___ (6) M___ 9
   Insight____ Efficiency score

2. Angle: Time: a2___ b1___ c1___ T___ 1/10___
   18-24-30 (6) Moves: a___ b___ c___ (9) M___ 21___1332
   Insight____ Efficiency score

3. Size-Shape Time: a2___ b1___ c1___ T___ 1/10___
   38-46-54 (8) Moves: a___ b___ c___ (12) M___ 36___1284
   Insight____ Efficiency score

4. Size Comb. Time: a2___ b1___ c1___ T___ 1/10___
   60-66-72 (6) Moves: a___ b___ c___ (9) M___ 48___1236
   Insight____ Efficiency score

5. Size-Shape Comb: Time: a4___ b1___ T___ 1/10___
   80-88 (8) Moves: a___ b___ (8) M___ 62___1188
   Insight____ Efficiency score

6. Rotation-Invers. Time: a4___ b1___ T___ 1/10___
   104-120 (16) Moves: a___ b___ (16) M___ 84___1140
   Insight____ Efficiency score

7. Creative Combin. Time: a8___ b1___ T___ 1/10___
   132-144 (12) Moves: a___ b___ (12) M___ 108___1080
   Insight____ Efficiency score

8. Rocker Function: Time: a8___ b1___ T___ 1/10___
   150-156 (6) Moves: a___ b___ (6) M___ 126___ 972
   Insight____ Efficiency score

9. Wedge Function: Time: a15__ b3___ T___ 1/10___
   164-172 (8) Moves: a___ b___ (8) M___ 154___ 864
   Insight____ Efficiency score

10. Creative Funct: Time: a15__ b3___ T___ 1/10___
    195-218 (23) Moves: a___ b___ (20) M___ 194___ 648
    Insight____ Efficiency score

11. Eccentric Funct: Time: a15__ b3___ T___ 1/10___
    232-246 (14) Moves: a___ b___ (14) M___ 228___ 432
    Insight____ Efficiency score

12. Adv. Creat. Funct: Time: a15__ b3___ T___ 1/10___
    264-282 (18) Moves: a___ b___ (18) M___ ___ 216
    Insight____ Efficiency score

TOTAL INSIGHT SCORE (ADD cumulative score of highest "too easy" test. ____

TOTAL EFFICIENCY SCORE (Note: Add cumulative score of highest "too easy" test--left hand figure -- and cumulative score of lowest "too hard" test--right hand figure--to total of scores made.) ____

KELLY SPATIAL INSIGHT TEST
Instructions for scoring and using score sheet

Most performance tests are comprised of tasks which require a comparatively low level of insight. Difficult tasks are devised merely by compounding tasks which an infant could accomplish if his attention span were great enough. This deficiency in performance testing has probably arisen out of the need for simple instructions which would circumvent verbal difficulties.

The present test is designed to sample the subject's ability to perform at various levels of insight. All of the tasks are in the field of spatial relations. A comprehensive measurement of a person's ability should also include tasks in other areas.

Each of the boards is named according to the type of insight required to solve it.

1. Gross Shape Discrimination
2. Angle--Discrimination of an attibute of shape
3. Size-Shape--Discrimination of space-filling fitness both in size and shape
4. Size Combination--Building areas to meet specifications
5. Size-Shape Combination--Building areas while disregarding minor shape cues.
6. Rotation-Inversion--Orientation
7. Creative Combination--Building areas without recess cues
8. Rocker Function--Devising an instrument to perform a functi
9. Wedge Function--Exploitation of forces where motions are no apparent
10. Creative Function--Use of wedge and rocker without recess c
11. Eccentric Function--Special form of the wedge and rotary mo
12. Advanced Creative Function--Eccentric and Rocker without recess cues

Each of the first four tests is administered three times and each of the remaining tests is administered twice (See Instructions for Administration). Record time in seconds for each trial and ente on the "Time" line after "a", "b", and "c" respectively. Record the total time for all three (or two) trials after "T". Take one tenth (nearest whole number) of this total time and record it after "1/10" Record the number of moves in each trial and record on the "moves" line after "a", "b", and "c" respectively. Record the total number of moves on all three trials after "M". Add this score to the figur immediately above it (one tenth of the time in seconds) to obtain the "Efficiency Score". Enter "Efficiency Score" for the test.

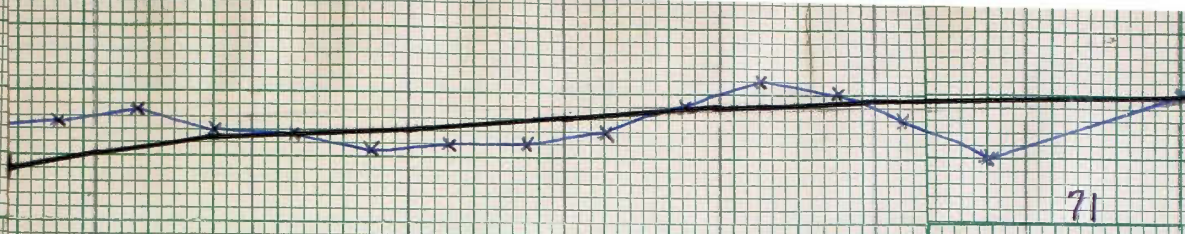
One minute before the time limit on the first trial of each test say, "You have just thirty seconds left. Make the best solutic you can and I will score it as it stands one minute from now." At the end of the <u>first trial only</u> score two points for each piece properly placed except Test 10 where the outside cross piece counts five points and Test 11 where the pivot nail counts four points. In Test 12 the two inferior solutions count 13 points. This score is the "Insight" score. Enter it after "Insight."

Do not give tests which are obviously too hard or too easy. Cumulative scores are given on the blank for inferring optimum and minimum performances on all tests below or above a certain test.

weighted insight scores

Fig. 1

Fig. 2



Combined weighted scores

Fig. 3

BIBLIOGRAPHY

1. Arthur, Grace. A Point Scale of Performance Tests. Clinical Manual. Vol. I. New York, The Commonwealth Fund Division of Publications, 1930. 82 p.

An explanation of the procedure in standardizing a point performance scale.

2. Arthur, Grace. The Process of Standardization. Vol. 2. New York, The Commonwealth Fund Division of Publications, 1933. 106 p.

Review literature in the field.

3. Baker, Winifred J. "Further Studies of Learning Aptitudes in Foreign-Born Prediction of Grade School Aptitude." (Trans. Kans. Acad. Sci., 1942. 5 p.)

Comparison of ground school grades and the measure of aptitude.

4. Cornell, Ethel L. and Coxe, Warren W. A Performance Ability Scale. Examination Manual. New York, The World Book Company, 1934. 68 p.

Description and explanation of the Cornell-Coxe performance scale.

BIBLIOGRAPHY

1. Arthur, Grace. A Point Scale of Performance Tests. Clinical Manual. Vol. 1. New York, The Commonwealth Fund Division of Publications, 1930. 82 p.

An explanation of the procedure in standardizing a point performance scale.

2. Arthur, Grace. The Process of Standardization. Vol. 2. New York, The Commonwealth Fund Division of Publications, 1933. 106 p.

Review literature in the field.

3. Baker, Winifred A. "Further Studies of Learning Aptitudes in Pilots--B: Prediction of Ground School Aptitude." (Trans. Kans. Acad. Sci., 1942. 5 p.)

Comparison of ground school grades and the measure of aptitude.

4. Cornell, Ethel L. and Coxe, Warren W. A Performance Ability Scale. Examination Manual. New York, The World Book Company, 1934. 88 p.

Description and explanation of the Cornell-Coxe performance scale.

5. Garrett, Henry E. Statistics in Psychology and Education. 2nd edition. New York, London, etc., Longmans, Green and Company, 1939. 493 p.

Statistical formulas used in plotting the norms.

6. Kent, Grace H. "Suggestions for the Next Revision of the Binet-Simon Scale." (in The Psychological Record. Vol. 1, No 25, November, 1937. pp. 409-433).

Other literature in the field.

7. Matheny, William Guy. "The Effectiveness of Aptitude Testing in the Civilian Pilot Training Program." (Trans. Kans. Acad. Sci., Vol. 44, 1941. pp. 363-65.)

One way the Kelly Spatial Insight Test has proved to be an effective test.

8. Matheny, William Guy. "Further Studies of Learning Aptitude in Pilots--A: Prediction of Flight Aptitude." (Trans. Kans. Acad. Sci., 1942. 4 p.)

Comparison of ratings given by flight operator and flight examiner.

9. Pintner, Rudolf and Paterson, Donald G. A Scale of Performance Tests. New York, London, D. Appleton and Company, 1931. 213 p.

Review of literature in the field.

10. Kelly, George A. "A New Form Board Scale." (In the Psychoclinician. Vol. 2, No. 2, June, 1937. pp. 40-43).

An explanation of the Kelly Spatial Insight Test.

11. Terman, Lewis M. The Measurement of Intelligence. An explanation of and a complete guidefor the use of the Stanford revision and extension of the Binet-Simon intelligence scale. Boston, New York, etc., Houghton Mifflin Company, The Riverside Press, Cambridge, 1916, 362 p.

First revision of the Binet test of mental ability.

12. Terman, Lewis M. and Merrill, Maud A. Measuring Intelligence. A guide to the administration of the new revised Standord-Binet tests of intelligence. Boston, New York, etc., Houghton Mifflin Company, The Riverside Press, Cambridge, 1937. 461 p.

A revised edition of the Stanford-Binet test.